

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:02:34 ; Search time 148 Seconds
(without alignments)
1672.821 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSWKCLFMAVLYATLCT.....KPDYPLFDGIFCSHHHHH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 segs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US50_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2063	100.0	384	19	US-09-500-376-8
2	2063	100.0	384	24	US-10-062-809-8
3	1922	93.2	361	15	US-09-175-683-12
4	1922	93.2	361	15	US-09-175-683B-10
5	1922	93.2	361	15	US-09-175-684A-10
6	1919.5	93.0	376	15	US-09-175-683-13

7	1919.5	93.0	379	15	US-09-175-683B-11	Sequence 11, Appl
8	1919.5	93.0	379	15	US-09-175-684A-11	Sequence 11, Appl
9	1879	91.1	378	24	PCT-US02-06415-11	Sequence 11, Appl
10	1879	91.1	378	24	US-10-087-464-11	Sequence 11, Appl
11	1879	91.1	400	21	US-09-791-537-54386	Sequence 54386, A
12	1879	91.1	539	21	US-09-791-537-107657	Sequence 107657, A
13	1879	91.1	539	21	US-09-791-537-107661	Sequence 107661, A
14	1879	91.1	539	21	US-09-791-537-108117	Sequence 108117, A
15	1879	91.1	539	21	US-09-791-537-108146	Sequence 108146, A
16	1879	91.1	613	15	US-08-593-006-22	Sequence 22, Appl
17	1879	91.1	613	15	US-09-117-415-22	Sequence 22, Appl
18	1879	91.1	613	15	US-09-117-415B-22	Sequence 22, Appl
19	1879	91.1	631	9	US-08-593-006-18	Sequence 18, Appl
20	1879	91.1	631	9	US-08-593-006-20	Sequence 20, Appl
21	1879	91.1	631	15	US-09-117-415-18	Sequence 18, Appl
22	1879	91.1	631	15	US-09-117-415-20	Sequence 20, Appl
23	1879	91.1	631	15	US-09-117-415B-18	Sequence 18, Appl
24	1879	91.1	631	15	US-09-117-415B-20	Sequence 20, Appl
25	1879	91.1	649	15	US-08-593-006-16	Sequence 16, Appl
26	1879	91.1	649	15	US-09-117-415-16	Sequence 16, Appl
27	1879	91.1	649	15	US-09-117-415B-16	Sequence 16, Appl
28	1879	91.1	1602	16	US-09-269-874A-7	Sequence 16, Appl
29	1879	91.1	1621	16	US-09-269-874A-5	Sequence 5, Appl
30	1879	91.1	1630	21	US-09-791-537-51240	Sequence 51240, A
31	1879	91.1	1639	16	PCT-US02-06415-10	Sequence 10, Appl
32	1879	91.1	1639	16	US-09-269-874-3	Sequence 3, Appl
33	1879	91.1	1639	21	US-09-269-874A-3	Sequence 3, Appl
34	1879	91.1	1639	21	US-09-791-537-51243	Sequence 51243, A
35	1879	91.1	1639	21	US-09-791-537-55902	Sequence 55902, A
36	1879	91.1	1639	24	US-10-087-464-10	Sequence 10, Appl
37	1876	90.9	539	21	US-09-791-537-107677	Sequence 107677, A
38	1876	90.9	539	21	US-09-791-537-108135	Sequence 108135, A
39	1874	90.8	355	15	US-09-175-683-11	Sequence 11, Appl
40	1874	90.8	355	15	US-09-175-683B-9	Sequence 9, Appl
41	1874	90.8	355	15	US-09-175-684A-9	Sequence 9, Appl
42	1874	90.8	376	1	PCT-US02-06415-34	Sequence 34, Appl
43	1874	90.8	376	17	US-09-311-817-2	Sequence 2, Appl
44	1874	90.8	376	23	US-09-978-756-2	Sequence 2, Appl
45	1874	90.8	376	24	US-10-087-464-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1

US-09-500-376-8

Sequence 8, Application US/09500376

GENERAL INFORMATION:

APPLICANT: University of Hawaii

TITLE OF INVENTION: Baculovirus Produced Plasmodium Falciiparum Vaccine

FILE REFERENCE: A-67984

CURRENT APPLICATION NUMBER: US/09/500, 376

CURRENT FILING DATE: 2000-02-08

PRIOR APPLICATION NUMBER: US 08/195, 705

PRIOR FILING DATE: 1994-02-14

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 8

LENGTH: 384

TYPE: PRT

ORGANISM: Plasmodium falciiparum

US-09-500-376-8

Query Match 100.0%; Score 2063; DB 19; Length 384;

Best Local Similarity 100.0%; Pred. No. 1.1e-160;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSWKCLFMAVLYATLCTAGAAVPSYIDNLSKENEYEVLYTKPLAGYRSLSKOL 60

DB 1 MMSWKCLFMAVLYATLCTAGAAVPSYIDNLSKENEYEVLYTKPLAGYRSLSKOL 60

QY 61 ENNVMTFNVNVKDIILSRFNKRENFNVLESIDIPYKDLTSSNYVVKDPRKLNKKRDK 120

Db 61 ENNVTEFVNVKDIINSFNRKRENFKNVLESOLIPYKDLTSSNVYKDPYFLNKEKDK 120
QY 121 FLSSTNYIKDSIDIDINFAVDVLYGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNN 180
Db 121 FLSSTNYIKDSIDIDINFAVDVLYGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNN 180
QY 181 IETLYKYVNHKIDLFVHLEAKVNLVYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFV 240
Db 181 IETLYKYVNHKIDLFVHLEAKVNLVYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFV 240
QY 241 GIADISTYNNHNNILTKFSTGWFENLAKTVLSNLIDGNLQGMNISOHOCVYKQCPON 300
Db 241 GIADISTYNNHNNILTKFSTGWFENLAKTVLSNLIDGNLQGMNISOHOCVYKQCPON 300
QY 301 SCGFHLDREBECKCLLNYKQEGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKIT 360
Db 301 SCGFHLDREBECKCLLNYKQEGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKIT 360
QY 361 CECTKPDSPYPLFDGIFCSHHHHH 384
Db 361 CECTKPDSPYPLFDGIFCSHHHHH 384

RESULT 2

US-10-062-809-8
; Sequence 8, Application US/10062809
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra P.
; APPLICANT: Hashimoto, Ann
; APPLICANT: Nishimura, Tani
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM FALCIPARUM VACCINE
; FILE REFERENCE: A-67984-1/RT/ITAL/NBC
; CURRENT APPLICATION NUMBER: US/10/062, 809
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/500, 376
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 60/266, 281
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-062-809-8.

Query Match 100.0%; Score 2063; DB 24; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.1e-160;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSWKCLLFMAVLVYATVCTAGAAVTPSVIDNIIISKIENEYEVLYKPLAGVYRSLSKROL 60
Db 1 MMSWKCLLFMAVLVYATVCTAGAAVTPSVIDNIIISKIENEYEVLYKPLAGVYRSLSKROL 60
QY 61 ENNVTEFVNVKDIINSFNRKRENFKNVLESOLIPYKDLTSSNVYKDPYFLNKEKDK 120
Db 61 ENNVTEFVNVKDIINSFNRKRENFKNVLESOLIPYKDLTSSNVYKDPYFLNKEKDK 120
QY 121 FLSSTNYIKDSIDIDINFAVDVLYGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNN 180
Db 121 FLSSTNYIKDSIDIDINFAVDVLYGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNN 180
QY 181 IETLYKYVNHKIDLFVHLEAKVNLVYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFV 240
Db 181 IETLYKYVNHKIDLFVHLEAKVNLVYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFV 240
QY 241 GIADISTYNNHNNILTKFSTGWFENLAKTVLSNLIDGNLQGMNISOHOCVYKQCPON 300
Db 241 GIADISTYNNHNNILTKFSTGWFENLAKTVLSNLIDGNLQGMNISOHOCVYKQCPON 300
QY 301 SCGFHLDREBECKCLLNYKQEGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKIT 360
Db 301 SCGFHLDREBECKCLLNYKQEGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKIT 360

QY 361 CECTKPDSPYPLFDGIFCSHHHHH 384
Db 361 CECTKPDSPYPLFDGIFCSHHHHH 384

RESULT 3

US-09-175-683-12
; Sequence 12, Application US/09175683A
; GENERAL INFORMATION:
; APPLICANT: Chen, Li How
; APPLICANT: Meade, Henry
; TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS
; TITLE OF INVENTION: FOR INCREASING MRNA LEVELS AND PROTEIN EXPRESSIONS IN
; FILE REFERENCE: 10275/134001
; CURRENT APPLICATION NUMBER: US/09/175, 683A
; CURRENT FILING DATE: 1998-10-20
; EARLIER APPLICATION NUMBER: US 60/095, 649
; EARLIER FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/062, 592
; EARLIER FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Preferably, a bacterium, virus, or parasite
US-09-175-683-12

Query Match 93.2%; Score 1922; DB 15; Length 361;
Best Local Similarity 99.7%; Pred. No. 3.9e-149;
Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 AVTBSVIDNIIISKIENEYEVLYKPLAGVYRSLSKOLENNVMTFVNVKDIINSFNRKRE 83
Db 1 AVTBSVIDNIIISKIENEYEVLYKPLAGVYRSLSKOLENNVMTFVNVKDIINSFNRKRE 60
QY 84 NFKVLESOLIPYKDLTSSNVYKDPYFLNKEKDKFLSSYNIKDSIDIDINFAVDV 143
Db 61 NFKVLESOLIPYKDLTSSNVYKDPYFLNKEKDKFLSSYNIKDSIDIDINFAVDV 120
QY 144 GYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNNIETLYKYVNHKIDLFVHLEAKV 203
Db 121 GYKILSEKYSKSDLSIKKYINDKOGENEKYLPFLNNIETLYKYVNHKIDLFVHLEAKV 180
QY 204 LNTYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFVGIADISTYNNHNNILTKFSTG 263
Db 181 LNTYTESKSNVEYKIKELNYLKTIDOKLADFKKNNNFVGIADISTYNNHNNILTKFSTG 240
QY 264 VFEINLAKTVLSNLIDGNLQGMNISOHOCVYKQCPONSGCFRHLDERBECKCLLNYKQEG 323
Db 241 VFEINLAKTVLSNLIDGNLQGMNISOHOCVYKQCPONSGCFRHLDERBECKCLLNYKQEG 300
QY 324 DKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKITCECTKPDSPYPLFDGIFCSHHHH 383
Db 301 DKCVENPPTCENNNGGCDADAKCTEEDSGSNGKKITCECTKPDSPYPLFDGIFCSHHHH 360
QY 384 H 384
Db 361 H 361

RESULT 4

US-09-175-683B-10
; Sequence 10, Application US/09175683B
; GENERAL INFORMATION:
; APPLICANT: Chen, Li How
; APPLICANT: Meade, Henry M.
; TITLE OF INVENTION: NOVEL MODIFIED NUCLEIC ACID SEQUENCES
; TITLE OF INVENTION: AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
; FILE REFERENCE: 10275-134001

; CURRENT APPLICATION NUMBER: US/09/175,683B
 ; CURRENT FILING DATE: 1998-10-20
 ; PRIOR APPLICATION NUMBER: US 60/085,649
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: US 60/062,592
 ; PRIOR FILING DATE: 1997-10-20
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 361
 ; TYPE: PRF
 ; ORGANISM: Plasmodium falciparum
 ; US-09-175-683B-10

Query Match 93.2%; Score 1922; DB 15; Length 361;
 Best Local Similarity 99.7%; Pred. No. 3.9e-149;
 Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 AVTPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNVNDILNSRPNKE 83
 DB 1 AVTPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNVNDILNSRPNKE 60
 QY 84 NFKNVLESDLIPYKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKDSIDTIDINPANDVL 143
 DB 61 NFKNVLESDLIPYKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKDSIDTIDINPANDVL 120
 QY 144 GYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVNHKIDLFVYIHLEAVY 203
 DB 121 GYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVNHKIDLFVYIHLEAVY 180
 QY 204 LNTYTERKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 263
 DB 181 LNTYTERKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 240
 QY 264 VFNENLAKTYLSMLDGNLQGMNLISQHCYVKQCPQNSGCFPHLDERECCKLLANYKQBG 323
 DB 241 VFNENLAKTYLSMLDGNLQGMNLISQHCYVKQCPQNSGCFPHLDERECCKLLANYKQBG 300
 QY 324 DKCVENPNTCENNGCGDADAKCTEEDSGSNGKKTCTCTPDSTPLFDGJFCSHHHH 383
 DB 301 DKCVENPNTCENNGCGDADAKCTEEDSGSNGKKTCTCTPDSTPLFDGJFCSHHHH 360
 QY 384 H 384
 DB 361 H 361

RESULT 5

; US-09-175-684A-10
 ; Sequence 10, Application US/09175684A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Li How
 ; APPLICANT: Meade, Harry M.
 ; TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID
 ; TITLE OF INVENTION: SEQUENCES AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
 ; FILE REFERENCE: 10275-133001
 ; CURRENT APPLICATION NUMBER: US/09/175,684A
 ; CURRENT FILING DATE: 1998-10-20
 ; PRIOR APPLICATION NUMBER: US 60/085,649
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: US 60/062,592
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 361
 ; TYPE: PRF
 ; ORGANISM: Plasmodium falciparum
 ; US-09-175-684A-10

Query Match 93.2%; Score 1922; DB 15; Length 361;
 Best Local Similarity 99.7%; Pred. No. 3.9e-149;

Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 24 AVTPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNVNDILNSRPNKE 83
 DB 1 AVTPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNVNDILNSRPNKE 60
 QY 84 NFKNVLESDLIPYKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKDSIDTIDINPANDVL 143
 DB 61 NFKNVLESDLIPYKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKDSIDTIDINPANDVL 120
 QY 144 GYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVNHKIDLFVYIHLEAVY 203
 DB 121 GYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVNHKIDLFVYIHLEAVY 180
 QY 204 LNTYTERKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 263
 DB 181 LNTYTERKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 240
 QY 264 VFNENLAKTYLSMLDGNLQGMNLISQHCYVKQCPQNSGCFPHLDERECCKLLANYKQBG 323
 DB 241 VFNENLAKTYLSMLDGNLQGMNLISQHCYVKQCPQNSGCFPHLDERECCKLLANYKQBG 300
 QY 324 DKCVENPNTCENNGCGDADAKCTEEDSGSNGKKTCTCTPDSTPLFDGJFCSHHHH 383
 DB 301 DKCVENPNTCENNGCGDADAKCTEEDSGSNGKKTCTCTPDSTPLFDGJFCSHHHH 360
 QY 384 H 384
 DB 361 H 361

RESULT 6

; US-09-175-683-13
 ; Sequence 13, Application US/09175683A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Li How
 ; APPLICANT: Meade, Henry
 ; TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS
 ; TITLE OF INVENTION: FOR INCREASING MRNA LEVELS AND PROTEIN EXPRESSIONS IN
 ; FILE REFERENCE: 10275/134001
 ; CURRENT APPLICATION NUMBER: US/09/175,683A
 ; CURRENT FILING DATE: 1998-10-20
 ; EARLIER APPLICATION NUMBER: US 60/095,649
 ; EARLIER FILING DATE: 1998-08-07
 ; EARLIER APPLICATION NUMBER: US 60/062,592
 ; EARLIER FILING DATE: 1997-10-20
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 376
 ; TYPE: PRF
 ; ORGANISM: preferably, a bacterium, virus, or parasite
 ; US-09-175-683-13

Query Match 93.0%; Score 1919.5; DB 15; Length 376;
 Best Local Similarity 97.3%; Pred. No. 6.6e-149;
 Matches 364; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 12 VLVATATL-CTAGAATPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNV 70
 DB 3 VLVATATL-CTAGAATPSVIDNLSIKIENEYEVLYLKPLAGYRSRLKQLENNVMTFNNV 62
 QY 71 VVDILNSRPNKRKNFNVLESOLIPKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKD 130
 DB 63 VVDILNSRPNKRKNFNVLESOLIPKDLTSSNYYVVKDPYKRLNKRKDFLSSYNYIKD 122
 QY 131 SIDTIDINPANDVLGYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVNH 190
 DB 123 SIDTIDINPANDVLGYKILSEKYSKSDLSIKKIYINDKQGENEKYLPFLNNITLYKYVND 182
 QY 191 KIDLFVYIHLEAVYLNNTYTERKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYN 250

```
Db 183 KIDFVHLEAKVLYQYITEKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIADLSTDYN 242
QY 251 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 310
Db 243 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 302
QY 311 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 370
Db 303 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 362
QY 371 LFDGIFCSHHNNH 384
Db 363 LFDGIFCSHHNNH 376

RESULT 7
US-09-175-683B-11
: Sequence 11, Application US/09175683B
: GENERAL INFORMATION:
: APPLICANT: Chen, Li How
: TITLE OF INVENTION: NOVEL MODIFIED NUCLEIC ACID SEQUENCES
: TITLE OF INVENTION: AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
: FILE REFERENCE: 10275-134001
: CURRENT APPLICATION NUMBER: US/09/175, 683B
: CURRENT FILING DATE: 1998-10-20
: PRIOR APPLICATION NUMBER: US 60/085,649
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: US 60/062,592
: PRIOR FILING DATE: 1997-10-20
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 379
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: altered MSP sequence; preferably, a bacterium,
: OTHER INFORMATION: virus, or parasite
US-09-175-683B-11

Query Match 93.08; Score 1919.5; DB 15; Length 379;
Best Local Similarity 97.38; Pred. No. 6.7e-149;
Matches 364; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 12 VLVATATL-CTAGAAVTPSVIDNIIISKIENEYEVLYLPLAGVYSRLKOLENNVMTFNVN 70
Db 3 VLIACIYALAAIAAVTPSVIDNIIISKIENEYEVLYLPLAGVYSRLKOLENNVMTFNVN 62
QY 71 VKDILNSRFNKRKNFKNVLESDDLIPYKDLTSSNVVADPYKFLNKRERDRLSSYNTIKD 130
Db 63 VKDILNSRFNKRKNFKNVLESDDLIPYKDLTSSNVVADPYKFLNKRERDRLSSYNTIKD 122
QY 131 SIDIDINFANDVLYGYKILSEKYSDDLISIKKYINDOGENEKKYLPPLNNIETLYKTVNH 190
Db 123 SIDIDINFANDVLYGYKILSEKYSDDLISIKKYINDOGENEKKYLPPLNNIETLYKTVND 182
QY 191 KIDFVHLEAKVLYQYITEKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIADLSTDYN 250
Db 183 KIDFVHLEAKVLYQYITEKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIADLSTDYN 242
QY 251 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 310
Db 243 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 302
QY 311 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 370
Db 303 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 362
QY 371 LFDGIFCSHHNNH 384
Db 363 LFDGIFCSHHNNH 376
```

```
RESULT 8
US-09-175-684A-11
: Sequence 11, Application US/09175684A
: GENERAL INFORMATION:
: APPLICANT: Chen, Li How
: TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID
: TITLE OF INVENTION: SEQUENCES AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
: FILE REFERENCE: 10275-133001
: CURRENT APPLICATION NUMBER: US/09/175, 684A
: CURRENT FILING DATE: 1998-10-20
: PRIOR APPLICATION NUMBER: US 60/085,649
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: US 60/062,592
: PRIOR FILING DATE: 1997-10-20
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 379
: TYPE: PRT
: ORGANISM: Plasmodium falciparum
US-09-175-684A-11

Query Match 93.08; Score 1919.5; DB 15; Length 379;
Best Local Similarity 97.38; Pred. No. 6.7e-149;
Matches 364; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 12 VLVATATL-CTAGAAVTPSVIDNIIISKIENEYEVLYLPLAGVYSRLKOLENNVMTFNVN 70
Db 3 VLIACIYALAAIAAVTPSVIDNIIISKIENEYEVLYLPLAGVYSRLKOLENNVMTFNVN 62
QY 71 VKDILNSRFNKRKNFKNVLESDDLIPYKDLTSSNVVADPYKFLNKRERDRLSSYNTIKD 130
Db 63 VKDILNSRFNKRKNFKNVLESDDLIPYKDLTSSNVVADPYKFLNKRERDRLSSYNTIKD 122
QY 131 SIDIDINFANDVLYGYKILSEKYSDDLISIKKYINDOGENEKKYLPPLNNIETLYKTVNH 190
Db 123 SIDIDINFANDVLYGYKILSEKYSDDLISIKKYINDOGENEKKYLPPLNNIETLYKTVND 182
QY 191 KIDFVHLEAKVLYQYITEKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIADLSTDYN 250
Db 183 KIDFVHLEAKVLYQYITEKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIADLSTDYN 242
QY 251 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 310
Db 243 HNNLLTFLSTGMEVFNIAKTVLSNLDGNLQGMINTISOHCYVKKOCPONSGCFRHLDER 302
QY 311 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 370
Db 303 EECCLNTYKOEKGCYENPPTCENNGGCDADAKTEEDSGSNGKRTICECTKPDSP 362
QY 371 LFDGIFCSHHNNH 384
Db 363 LFDGIFCSHHNNH 376

RESULT 9
PCT-US02-06415-11
: Sequence 11, Application PC/TUS0206415
: GENERAL INFORMATION:
: APPLICANT: St. Elizabeth's Medical Center, Inc.
: TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses There
: FILE REFERENCE: 51237/77019WC
: CURRENT APPLICATION NUMBER: PCT/US02/06415
: CURRENT FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 06/272,930
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 59
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 11
```

LENGTH: 378
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 PCT-US02-06415-11

Query Match 91.1%; Score 1879; DB 1; Length 378;
 Best Local Similarity 99.4%; Pred. No. 1,4e-145;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

22 GAAVPSVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 81
 1 GEAVTSPVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 60
 82 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 141
 61 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 120
 142 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 201
 121 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 180
 202 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 261
 181 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 240
 262 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 321
 241 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 300
 322 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 378
 301 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 357

RESULT 10
 US-10-087-464-11
 Sequence 11, Application US/10087464

GENERAL INFORMATION:
 APPLICANT: Chishell, Athar
 APPLICANT: Oh, Steven
 APPLICANT: Liu, David
 APPLICANT: Goel, Vikas
 APPLICANT: Li, Xuerong
 TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
 FILE REFERENCE: S1237/7019
 CURRENT APPLICATION NUMBER: US/10/087,464
 CURRENT FILING DATE: 2002-03-01
 PRIOR APPLICATION NUMBER: US 06/272,930
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 59
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 11
 LENGTH: 378
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-10-087-464-11

Query Match 91.1%; Score 1879; DB 24; Length 378;
 Best Local Similarity 99.4%; Pred. No. 1,4e-145;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

22 GAAVPSVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 81
 1 GEAVTSPVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 60
 82 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 141
 61 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 120
 142 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 201
 121 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 180

202 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 261
 181 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 240
 262 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 321
 241 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 300
 322 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 378
 301 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 357

RESULT 11
 US-09-791-537-54386
 Sequence 54386, Application US/09791537

GENERAL INFORMATION:
 APPLICANT: Biomolix, Inc.
 APPLICANT: Debe, Derek
 APPLICANT: Danzer, Joseph
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537
 CURRENT FILING DATE: 2001-02-22
 NUMBER OF SEQ ID NOS: 153055
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 54386
 LENGTH: 400
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-09-791-537-54386

Query Match 91.1%; Score 1879; DB 21; Length 400;
 Best Local Similarity 99.4%; Pred. No. 1,6e-145;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

22 GAAVPSVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 81
 23 GEAVTSPVIDNLSIKIENEYEVLYKPLAGVRSLSKKOLENNVTFNVNKKIDILNSRPNK 82
 82 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 141
 83 RENFKNVLESDDLPRYDLTSSNVVDPYKFLNKKERDRKFLSSYNIKSIDIDINPAND 142
 142 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 201
 143 VLGYRILSEKYSKSDLSIKKYINDKQGENEKLPLPLNIETLYKTVNKHIDLFVHLEA 202
 202 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 261
 203 KVLNTYKSNVEVKIKELNYLKTIDOKLADFKNNNFVGIALDSTDYNNHNLITFELST 262
 262 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 321
 263 GWVFENLAKTVLSNLDGNLQGMNLNISOHCYKQCPONSQCFRHLDERBECKCLNRYKQ 322
 322 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 378
 323 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTICECTKPDSPYPLFDGIFCS 379

RESULT 12
 US-09-791-537-107657
 Sequence 107657, Application US/09791537

GENERAL INFORMATION:
 APPLICANT: Biomolix, Inc.
 APPLICANT: Debe, Derek
 APPLICANT: Danzer, Joseph
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537

;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 107657
;; LENGTH: 539
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-09-791-537-107657

Query Match 91.1%; Score 1879; DB 21; Length 539;
Best Local Similarity 99.4%; Pred. No. 2.4e-145;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 81
DB 162 GEATVPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 221
QY 82 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 141
DB 222 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 281
QY 142 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 201
DB 282 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 341
QY 202 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 261
DB 342 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 401
QY 262 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 321
DB 402 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 461
QY 322 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 378
DB 462 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 518

RESULT 13

US-09-791-537-107661
;; Sequence 107661, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 107661
;; LENGTH: 539
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-09-791-537-107661

Query Match 91.1%; Score 1879; DB 21; Length 539;
Best Local Similarity 99.4%; Pred. No. 2.4e-145;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 81
DB 162 GEATVPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 221
QY 82 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 141
DB 222 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 281
QY 142 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 201
DB 282 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 341

QY 202 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 261
DB 342 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 401
QY 262 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 321
DB 402 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 461
QY 322 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 378
DB 462 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 518

RESULT 14

US-09-791-537-108117
;; Sequence 108117, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 108117
;; LENGTH: 539
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-09-791-537-108117

Query Match 91.1%; Score 1879; DB 21; Length 539;
Best Local Similarity 99.4%; Pred. No. 2.4e-145;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 81
DB 162 GEATVPSYIDNLSKIEEYEVLYKPLAGYRSIKKOLENNVMTFNVNVDIINSRPNK 221
QY 82 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 141
DB 222 RENKNVLESPLIPKDLTSSNYYVKDPYKFLNKKRKPFLSSYNYIKDSIDTDINPAND 281
QY 142 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 201
DB 282 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKTVNKKIDLFVYIHLEA 341
QY 202 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 261
DB 342 KLVNTYKSNVEYKIKELANTLKTIOKDLADFKKNNNFVGIAIDSTDYNNHNNLTJFLST 401
QY 262 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 321
DB 402 GMVEENLAKTVLSNLDGNLOGMLNISOHCVKQCPONSGCFRHLDERECKCLLANYKQ 461
QY 322 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 378
DB 462 EGDCKVENPNTCENNNGCGADAKCTEEDSGSNGKRTTCCTKPDSTPLPDGIFCS 518

RESULT 15

US-09-791-537-108146
;; Sequence 108146, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210

;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: Patent version 3.0
;; SEQ ID NO 108146
;; LENGTH: 539
;; TYPE: PRF
;; ORGANISM: Plasmodium falciparum
US-09-791-537-108146

Query Match 91.1%; Score 1879; DB 21; Length 539;
Best Local Similarity 99.4%; Pred. No. 2.4e-145;

Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	22	GAATPSYIDNLSKIENEYLYLPLAGYRSLSKOLENNVMTFNVVVKDILNSRFNK	81
DB	162	GEATPSYIDNLSKIENEYLYLPLAGYRSLSKOLENNVMTFNVVVKDILNSRFNK	221
QY	82	RENFKNVLESDLIPYKDLTSSNVVVKPYKFLNKKERDKFLSSYNTIKDSIDFTDINFAND	141
DB	222	RENFKNVLESDLIPYKDLTSSNVVVKPYKFLNKKERDKFLSSYNTIKDSIDFTDINFAND	281
QY	142	VLGYRKILSEKYSKSDLSIKKTIINDKQGENEKYLPPLNNIEPLTYKTVNKHIDLFVTHLEA	201
DB	282	VLGYRKILSEKYSKSDLSIKKTIINDKQGENEKYLPPLNNIEPLTYKTVNKHIDLFVTHLEA	341
QY	202	KVLNTYSEKSNVEVKEIKELNYLKTIDOKLADFFKNNNFYGIADLSTDYNNHNLTFKFLST	261
DB	342	KVLNTYSEKSNVEVKEIKELNYLKTIDOKLADFFKNNNFYGIADLSTDYNNHNLTFKFLST	401
QY	262	GAVFENLAKTVLSNLDGNLQGMNLISQHCYVKQCPONGSGCFRHLDERECKCLNTYKQ	321
DB	402	GAVFENLAKTVLSNLDGNLQGMNLISQHCYVKQCPONGSGCFRHLDERECKCLNTYKQ	461
QY	322	EGDKCVENPNPTCNENNGGCDADAKTEEDSGSNGKKTCECTKPDSPYPLFDGIFCS	378
DB	462	EGDKCVENPNPTCNENNGGCDADAKTEEDSGSNGKKTCECTKPDSPYPLFDGIFCS	518

Search completed: June 11, 2003, 16:09:14
Job time : 150 secs

THIS PAGE BLANK

Thu Jun 12 08:47:25 2003

us-09-500-376-8.rapn

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 11, 2003, 16:04:40 ; Search time 99 Seconds

(without alignments)
879,958 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063

Sequence: 1 MWSMKCLFMAVIVTATICT.....KPDSTPLDGFCSHHHHH 384

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1050967 seqs, 226864292 residues

Total number of hits satisfying chosen parameters: 1050967

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152.5	7.4	2184	6	US-10-304-095-6
2	149	7.2	688	7	US-60-453-135-12270
3	149	7.2	688	7	US-60-453-050-12270
4	149	7.2	688	7	US-60-455-444-6419
5	149	7.2	688	7	US-60-455-412-12270
6	149	7.2	688	7	US-60-455-412-12270
7	144.5	7.0	460	7	US-60-453-135-12271
8	144.5	7.0	460	7	US-60-453-050-12271
9	144.5	7.0	460	7	US-60-455-444-6420
10	144.5	7.0	460	7	US-60-455-444-6420
11	144.5	7.0	460	7	US-60-466-412-12271
12	144.5	7.0	802	7	US-60-453-135-12273
13	144.5	7.0	802	7	US-60-453-050-12273
14	144.5	7.0	802	7	US-60-455-444-6422
15	144.5	7.0	802	7	US-60-455-444-6422
16	144.5	7.0	802	7	US-60-466-412-12273
17	142.5	6.9	820	6	US-10-302-812-42
18	142	6.9	752	6	US-10-302-812-42
19	142	6.9	849	6	US-10-335-977-7264
20	142	6.9	849	6	US-10-335-977-7265
21	141.5	6.9	365	7	US-60-453-135-12267
22	141.5	6.9	365	7	US-60-453-050-12267
23	141.5	6.9	365	7	US-60-455-444-6416
24	141.5	6.9	365	7	US-60-466-412-12267
25	141.5	6.9	384	6	US-10-092-411A-4669
26	141.5	6.9	384	6	US-10-092-411A-4669

27	141.5	6.9	432	7	US-60-453-135-12268	Sequence 12268, A
28	141.5	6.9	432	7	US-60-453-050-12268	Sequence 12268, A
29	141.5	6.9	432	7	US-60-455-444-6417	Sequence 6417, Ap
30	141.5	6.9	432	7	US-60-466-412-12268	Sequence 12268, A
31	141.5	6.9	432	7	US-60-455-444-6417	Sequence 6417, Ap
32	141.5	6.9	609	5	US-09-949-016-7747	Sequence 7747, Ap
33	141.5	6.9	609	5	US-09-949-016-7748	Sequence 7748, Ap
34	141.5	6.9	609	5	US-09-949-016-7749	Sequence 7749, Ap
35	141.5	6.9	609	5	US-09-949-016-7750	Sequence 7750, Ap
36	141.5	6.9	609	5	US-09-949-016-7751	Sequence 7751, Ap
37	141.5	6.9	609	5	US-09-949-016-7752	Sequence 7752, Ap
38	141.5	6.9	609	5	US-09-949-016-7753	Sequence 7753, Ap
39	141.5	6.9	609	5	US-09-949-016-7754	Sequence 7754, Ap
40	141.5	6.9	702	5	US-09-724-676-82867	Sequence 82867, A
41	141.5	6.9	702	5	US-09-724-676-82868	Sequence 82868, A
42	141.5	6.9	702	5	US-09-724-676-82869	Sequence 82869, A
43	141.5	6.9	702	5	US-09-724-676-82870	Sequence 82870, A
44	141.5	6.9	702	5	US-09-724-676-82871	Sequence 82871, A
45	141.5	6.9	702	5	US-09-724-676-82872	Sequence 82872, A

ALIGNMENTS

RESULT 1
US-10-304-095-6
Sequence 6, Application US/10304095
GENERAL INFORMATION:
APPLICANT: Long, David M.
APPLICANT: Metz, Anneke M.
APPLICANT: Love, Ruschelle A.
TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
FILE REFERENCE: 47714-5009-US
CURRENT APPLICATION NUMBER: US/10/304,095
PRIOR FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: US/09/417,485
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2184
TYPE: PRT
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: unsure
LOCATION: (330)..
OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-10-304-095-6

Query Match
Best Local Similarity 7.4%; Score 152.5; DB 6; Length 2184;
Matches 93; Conservatively 54; Mismatches 133; Indels 139; Gaps 20;

25 VTPSIVNIIISKIENEVYLYLKLAVGVRSLKOLENNVMT-FNVNVD----- 73
125 ILKFFPNT-----NNEFTYL--FKRSFLYKRNKNNSEFLSYPPNVAVSFGLHTRNR 1307
74 --IINSEFKREN-----FKNYESDLIPYKDLTSSNVVADPKFLKRRDFTLS-- 123
1308 ELIKSHLNDNHFFLNQMKTKSKDL-----YFADSKSLQVVRKDFMTII 1357
124 ---SYNI-----KSDITDINPANDVLYGYYKILSEYKSDLSIRK-YIND 166
1358 TVIRIYLLNYSIKFEKFNKRNIYFYFOENQMKGYV--LSVDRKRVENIKKWLNS 1415
167 -KQGENEKYLPFLNNIETLYKTVNKRIDFLVHLFAKLVNTYKSNVYKIKELNLT 225
1416 MKKINDELIESLKNSI--NINNKNMICNHEQD---TEKGNTOXNEKHDIYICP 1468
226 IO---DKLADKKRNNVPG-IADISTDY----- 249
1469 IYNSPDSITTTTHSSNNYKGNNIHVSGDYKRDGLHKGNNSMECYVKDICKNNNNNN 1528

Db 219 ERSFHRPILOAGLPANKTVALGNSVEFMCKYSDPOPHIOMLKHIEVNGSKIGPDLPIY 278
QY 241 GI-----ADLSTDYNNHNLTK 257
Db 279 QILKVSADSSASMSNGVLLVR 299

RESULT 5

US-60-465-241-6419
; Sequence 6419, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BECOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: RHETIMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6419
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-6419

Query Match 7.2%; Score 149; DB 7; Length 688;
Best Local Similarity 25.9%; Pred. No. 1.3e-05;
Matches 83; Conservative 45; Mismatches 107; Indels 86; Gaps 18;

QY 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
Db 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
QY 60 LENNVMTFN-----VNVKDLNSRFNKRKNVLESGLIP-----YKDLTSS----- 102
Db 57 LRDDVOSINMLRDGVQLAESNRTITGEE-----VEQDSVPADSGIACVTSPPSGSDPT 112
QY 103 ----NYVADPYKFLNKRKDFL-----SSNYIKDSIDTDINF 138
Db 113 YFSVNSAVAPY-WTSPKMEKKLHAPAAKTVKFCPSGTPNPTLRMLKNKE-----F 167
QY 139 AND-VLGYYKILSEKSDLSIKKYINDKGENEKYLPFLNNIETLYKTVNKHIDLFYI 197
Db 168 KPDHRIGGYKVRATWSIIMDSV---VPSDKGN-----YTCIYENEGSINHRYOLDVY 218
QY 198 H-----LEAKV-LNTYVE-KSNVEVKIK-----ELNYLKTIDKLADEFKKN-NFY 240
Db 219 ERSFHRPILOAGLPANKTVALGNSVEFMCKYSDPOPHIOMLKHIEVNGSKIGPDLPIY 278
QY 241 GI-----ADLSTDYNNHNLTK 257
Db 279 QILKVSADSSASMSNGVLLVR 299

RESULT 6

US-60-466-412-12270
; Sequence 12270, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12270
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens

US-60-466-412-12270

Query Match 7.2%; Score 149; DB 7; Length 688;
Best Local Similarity 25.9%; Pred. No. 1.3e-05;
Matches 83; Conservative 45; Mismatches 107; Indels 86; Gaps 18;

QY 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
Db 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
QY 60 LENNVMTFN-----VNVKDLNSRFNKRKNVLESGLIP-----YKDLTSS----- 102
Db 57 LRDDVOSINMLRDGVQLAESNRTITGEE-----VEQDSVPADSGIACVTSPPSGSDPT 112
QY 103 ----NYVADPYKFLNKRKDFL-----SSNYIKDSIDTDINF 138
Db 113 YFSVNSAVAPY-WTSPKMEKKLHAPAAKTVKFCPSGTPNPTLRMLKNKE-----F 167
QY 139 AND-VLGYYKILSEKSDLSIKKYINDKGENEKYLPFLNNIETLYKTVNKHIDLFYI 197
Db 168 KPDHRIGGYKVRATWSIIMDSV---VPSDKGN-----YTCIYENEGSINHRYOLDVY 218
QY 198 H-----LEAKV-LNTYVE-KSNVEVKIK-----ELNYLKTIDKLADEFKKN-NFY 240
Db 219 ERSFHRPILOAGLPANKTVALGNSVEFMCKYSDPOPHIOMLKHIEVNGSKIGPDLPIY 278
QY 241 GI-----ADLSTDYNNHNLTK 257
Db 279 QILKVSADSSASMSNGVLLVR 299

RESULT 7

US-60-453-135-12271
; Sequence 12271, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12271
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-12271

Query Match 7.0%; Score 144.5; DB 7; Length 460;
Best Local Similarity 25.9%; Pred. No. 2e-05;
Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

QY 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
Db 1 MMSKCLLFMAVLATATCTAGAAVT-PSVIDNLSKIENEVEYLTKPLAGYRSIKKQ 59
QY 60 LENNVMTFN-----VNVKDLNSRFNKRKNVLESGLIP-----YKDLTSS----- 102
Db 57 LRDDVOSINMLRDGVQLAESNRTITGEE-----VEQDSVPADSGIACVTSPPSGSDPT 112
QY 103 ----NYVADPYKFLNKRKDFL-----SSNYIKDSIDTDINF 138
Db 113 YFSVNSAVAPY-WTSPKMEKKLHAPAAKTVKFCPSGTPNPTLRMLKNKE-----F 167
QY 139 AND-VLGYYKILSEKSDLSIKKYINDKGENEKYLPFLNNIETLYKTVNKHIDLFYI 197
Db 168 KPDHRIGGYKVRATWSIIMDSV---VPSDKGN-----YTCIYENEGSINHRYOLDVY 218
QY 198 H-----LEAKV-LNTYVE-KSNVEVKIK-----ELNYLKTIDKLADEFKKN-NFY 240
Db 219 ERSFHRPILOAGLPANKTVALGNSVEFMCKYSDPOPHIOMLKHIEVNGSKIGPDLPIY 278

RESULT 8

US-60-453-050-12271

Sequence 12271, Application US/60453050

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: LUK, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001457

CURRENT APPLICATION NUMBER: US/60/453,050

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82/62

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 12271

LENGTH: 460

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-050-12271

Query Match 7.0%; Score 144.5; DB 7; Length 460;

Best Local Similarity 25.9%; Pred. No. 2e-05; 90; Indels 81; Gaps 16;

Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 59

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

RESULT 10

US-60-465-241-6420

Sequence 6420, Application US/60465241

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: BEGOVICH, Ann

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001468

CURRENT APPLICATION NUMBER: US/60/465,241

CURRENT FILING DATE: 2003-04-23

NUMBER OF SEQ ID NOS: 258418

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 6420

LENGTH: 460

TYPE: PRT

ORGANISM: Homo sapiens

US-60-465-241-6420

Query Match 7.0%; Score 144.5; DB 7; Length 460;

Best Local Similarity 25.9%; Pred. No. 2e-05; 90; Indels 81; Gaps 16;

Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 59

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

1 MMSMKCLFMAVAVLTATCTAGAAVT-PSVIDNLSKIENEYEVLYLKLPLAGYRSLSKKQ 56

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12271
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-60-466-412-12271

Query Match
Best Local Similarity 7.0%; Score 144.5; DB 7; Length 460;
Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSKCLIFMAVLTATCTAGAAVT-PSVIDNLSIKENEYEVLYLKLPLAGYRSLSKQ 59
|||||
1 MMSKCLIFMAVLTATCTARPSPTLPDQAPGAPVEVSEFLVH---PGDILQLRCR 56
60 LENNVMTFN---VNVKDLNSRFNKRNFKNVLESIDLIP-----YKDLTSS----- 102
57 LRDDVOSIMWLRDGVQLAESNRTRITGEE---VEQDVSVPADSGIACVTSPPSSGSDTT 112
103 ---NVVADPYKFLNKKERKDFL-----SSYNTIKDSIDTDINF 138
113 YFSVNSAVALPY-WTSPKMERKLAHPAAKTVKFCPPSGTPTPLRLMKNGKE---F 167
139 AND-VLGYKILSEKYSKSDLSIKKYINDKGENEKYLPNNIETLYTVNKHIDLFVI 197
168 KPDHRIGYKVRATWISIMDSV---VPSDKGN-----YTCIVENETGSIHNTYQLDVY 218
198 H-----LEAKV-LNTYVE-KSNVEVKIR-----ELNVLKTIQ 227
219 ERSPHRPILOAGLPANKTVALGSNVEFMCKVYSDPQHIOMLKHIE 264

RESULT 12
US-60-453-135-12273

Sequence 12273, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAROUOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001456
CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12273
LENGTH: 802
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-135-12273

Query Match
Best Local Similarity 7.0%; Score 144.5; DB 7; Length 802;
Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSKCLIFMAVLTATCTAGAAVT-PSVIDNLSIKENEYEVLYLKLPLAGYRSLSKQ 59
|||||
1 MMSKCLIFMAVLTATCTARPSPTLPDQAPGAPVEVSEFLVH---PGDILQLRCR 56
60 LENNVMTFN---VNVKDLNSRFNKRNFKNVLESIDLIP-----YKDLTSS----- 102
57 LRDDVOSIMWLRDGVQLAESNRTRITGEE---VEQDVSVPADSGIACVTSPPSSGSDTT 112
103 ---NVVADPYKFLNKKERKDFL-----SSYNTIKDSIDTDINF 138
113 YFSVNSAVALPY-WTSPKMERKLAHPAAKTVKFCPPSGTPTPLRLMKNGKE---F 167
139 AND-VLGYKILSEKYSKSDLSIKKYINDKGENEKYLPNNIETLYTVNKHIDLFVI 197
168 KPDHRIGYKVRATWISIMDSV---VPSDKGN-----YTCIVENETGSIHNTYQLDVY 218
198 H-----LEAKV-LNTYVE-KSNVEVKIR-----ELNVLKTIQ 227

Db 219 ERSPHRPILOAGLPANKTVALGSNVEFMCKVYSDPQHIOMLKHIE 264

RESULT 13
US-60-453-050-12273
Sequence 12273, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
STENOSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12273
LENGTH: 802
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-050-12273

Query Match
Best Local Similarity 7.0%; Score 144.5; DB 7; Length 802;
Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSKCLIFMAVLTATCTAGAAVT-PSVIDNLSIKENEYEVLYLKLPLAGYRSLSKQ 59
|||||
1 MMSKCLIFMAVLTATCTARPSPTLPDQAPGAPVEVSEFLVH---PGDILQLRCR 56
60 LENNVMTFN---VNVKDLNSRFNKRNFKNVLESIDLIP-----YKDLTSS----- 102
57 LRDDVOSIMWLRDGVQLAESNRTRITGEE---VEQDVSVPADSGIACVTSPPSSGSDTT 112
103 ---NVVADPYKFLNKKERKDFL-----SSYNTIKDSIDTDINF 138
113 YFSVNSAVALPY-WTSPKMERKLAHPAAKTVKFCPPSGTPTPLRLMKNGKE---F 167
139 AND-VLGYKILSEKYSKSDLSIKKYINDKGENEKYLPNNIETLYTVNKHIDLFVI 197
168 KPDHRIGYKVRATWISIMDSV---VPSDKGN-----YTCIVENETGSIHNTYQLDVY 218
198 H-----LEAKV-LNTYVE-KSNVEVKIR-----ELNVLKTIQ 227
219 ERSPHRPILOAGLPANKTVALGSNVEFMCKVYSDPQHIOMLKHIE 264

RESULT 14
US-60-455-444-6422

Sequence 6422, Application US/60455444
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: BEGOVICH, Ann
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
RHUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001455
CURRENT APPLICATION NUMBER: US/60/455,444
CURRENT FILING DATE: 2003-03-18
NUMBER OF SEQ ID NOS: 50986
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6422
LENGTH: 802
TYPE: PRT
ORGANISM: Homo sapiens
US-60-455-444-6422

Query Match
Best Local Similarity 7.0%; Score 144.5; DB 7; Length 802;
Matches 74; Conservative 41; Mismatches 90; Indels 81; Gaps 16;

1 MMSKCLIFMAVLTATCTAGAAVT-PSVIDNLSIKENEYEVLYLKLPLAGYRSLSKQ 59
|||||
1 MMSKCLIFMAVLTATCTARPSPTLPDQAPGAPVEVSEFLVH---PGDILQLRCR 56

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:02:30 ; Search time 73 Seconds
(without alignments)
700.935 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSMKCLEFWAVLTATICT.....KPDSPLEFDGIFCSHHNNH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1922	93.2	361	20	AAV09373
2	1922	93.2	361	20	AAV05833
3	1919.5	93.0	376	20	AAV09374
4	1919.5	93.0	376	20	AAV05834
5	1879	91.1	1639	19	AAW54145
6	1874	90.8	355	20	AAV09372
7	1874	90.8	355	20	AAV05832
8	1684.5	81.7	1654	6	AAV0777
9	1326	64.3	262	22	AAV37610
10	1161.5	56.3	375	22	AAV83926

11	546	26.5	108	22	AAV37609
12	542	26.3	96	22	AAV37608
13	539	26.1	116	18	AAV36103
14	539	26.1	116	18	AAV22592
15	539	26.1	127	18	AAV22593
16	539	26.1	127	18	AAV36102
17	370.5	18.0	206	22	AAV00659
18	293	14.2	54	14	AAV41356
19	290	14.1	108	22	AAV47487
20	284.5	13.8	187	22	AAV85697
21	284.5	13.8	2028	22	AAV85698
22	282	13.7	49	14	AAV41354
23	279	13.5	49	14	AAV41355
24	274	13.3	54	14	AAV41357
25	268	13.0	93	22	AAV37611
26	253.5	12.3	106	14	AAV41358
27	196.5	9.5	350	21	AAV70278
28	168.5	8.2	980	21	AAV18294
29	152.5	7.4	2184	22	AAE00425
30	149	7.2	1714	21	AAV18275
31	148.5	7.2	2013	21	AAV18265
32	146.5	7.1	2539	21	AAV18198
33	146	7.1	622	21	AAV97170
34	144	7.0	816	14	AAV39647
35	143.5	7.0	722	21	AAV18291
36	143.5	7.0	1516	21	AAV18195
37	143.5	7.0	1712	21	AAV18205
38	143.5	7.0	2500	21	AAV18272
39	142.5	6.9	820	20	AAV06458
40	141.5	6.9	384	23	ABP39824
41	141.5	6.9	820	19	AAV63844
42	141.5	6.9	822	12	AAV13549
43	141.5	6.9	822	13	AAV20750
44	141.5	6.9	822	13	AAV26337
45	139.5	6.8	822	22	AAV04694

ALIGNMENTS

RESULT 1
AAV09373
ID AAV09373 standard; Protein; 361 AA.
XX
AC AAV09373;
XX
DT 31-ANG-1999 (first entry)
XX
DE Merozoite surface protein MSP-1-42.
XX
KW MSP-1: merozoite surface protein; malaria; vaccine;
KW protein engineering; protein expression; codon usage;
KW transgenic animal.
XX
OS Plasmodium falciparum.
PN WO920774-A2.
XX
PD 29-APR-1999.
XX
PF 20-OCT-1998; 98WO-US22226.
XX
PR 15-MAY-1998; 98US-0085649.
XX
PR 20-OCT-1997; 97US-0062592.
XX
PA (GENZ) GENZYME TRANSGENICS CORP.
XX
PI Chen LH, Meade H;
XX
DR WPI: 1999-288313/24.
DR F-PSDB; AAV56009.
XX
PT Modified malarial protein for use in anti-malarial vaccines

Query	Match	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
Qy	24	AVTPSVIDNIIISKIENEYEVLYKRLAGYRSIKRQLENNVTENVVAVDILNSRPNKRE	93.2%	1922;	361;	360;	0;	1;	0;	0
Db	1	AVTPSVIDNIIISKIENEYEVLYKRLAGYRSIKRQLENNVTENVVAVDILNSRPNKRE	99.7%	Pred. No. 3.5e-127;						
Qy	84	NFKNVLESDLIPKRLDTSSNYYVKDPYKFLNKKRDKFLSSYNTIKDSIDTIDINFANDVL								
Db	61	NFKNVLESDLIPKRLDTSSNYYVKDPYKFLNKKRDKFLSSYNTIKDSIDTIDINFANDVL								
Qy	144	GYTKILSKRYSDSLDIKKIKYINDKQGENEYKLPFLNNIETLYKTYNHRKIDLVHILEAVY								
Db	121	GYTKILSKRYSDSLDIKKIKYINDKQGENEYKLPFLNNIETLYKTYNHRKIDLVHILEAVY								
Qy	204	LNITYEKSNEVEKIKELNYLKTIDOKLADFKNNNEVGIADLSTVDYNNHNLTKFLSTGM								
Db	181	LNITYEKSNEVEKIKELNYLKTIDOKLADFKNNNEVGIADLSTVDYNNHNLTKFLSTGM								
Qy	264	VFEINAKTVLSNLDGNIQGMINTISOHCYVKQCPQNSGCFPHLDERECCKLLNYKQSG								
Db	241	VFEINAKTVLSNLDGNIQGMINTISOHCYVKQCPQNSGCFPHLDERECCKLLNYKQSG								
Qy	324	DKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTPDSYPLFDGIFCSHHNNH								
Db	301	DKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTPDSYPLFDGIFCSHHNNH								
Qy	384	H 384								
Db	361	H 361								

Query Match	Best Local Similarity	Score	DB	Length
Matches 360; Conservative	93.7%;	1922;	DR 20;	361;
	0;	Mismatches	1;	Indels
			0;	Gaps
				0
Sequence	361 AA;			
Query	24	AAVPSVIDNLSKIENEYEVLYLKLPLAGVYSLSKQLENNVTFFVNVKYLINSRPNKRE	83	
Db	1	AAVPSVIDNLSKIENEYEVLYLKLPLAGVYSLSKQLENNVTFFVNVKYLINSRPNKRE	60	
QY	84	NFKNVLSDDLIPYKDLISSTNVKDPYFLNKEKKDKFLSSYNTIKDSIDIDINFANDYL	143	
Db	61	NFKNVLSDDLIPYKDLISSTNVKDPYFLNKEKKDKFLSSYNTIKDSIDIDINFANDYL	120	
QY	144	GYVILSEKYSQSDLSIKKYLINDKQSENEKYLPFLNNIETLKYTVNHHKIDFLVHLEAKV	203	
Db	121	GYVILSEKYSQSDLSIKKYLINDKQSENEKYLPFLNNIETLKYTVNHHKIDFLVHLEAKV	180	
QY	204	LNTYERSNVEVRIKELNLYLTKTIDOKLADFKRNNNFVGIADLSTVDYNNHNLTKFLSTGM	263	
Db	181	LNTYERSNVEVRIKELNLYLTKTIDOKLADFKRNNNFVGIADLSTVDYNNHNLTKFLSTGM	240	
QY	264	VFENMLARTVLSNLDGNTQGMNLNSHQCVKQCKQNSGCGRRHLDREDECKSLMYKQBG	323	
Db	241	VFENMLARTVLSNLDGNTQGMNLNSHQCVKQCKQNSGCGRRHLDREDECKSLMYKQBG	300	
QY	324	DKCEVNPPTCENNNNGGCDADAKCEBDSGSGSKRIITPCETKRPDSYPLFDGIFGSHNNH	383	
Db	301	DKCEVNPPTCENNNNGGCDADAKCEBDSGSGSKRIITPCETKRPDSYPLFDGIFGSHNNH	360	
QY	384	H 384		
Db	361	H 361		


```

XX AC AAY09374;
XX 31-AUG-1999 (first entry)
XX DE Modified merozoite surface protein MSP-1-42.
XX KW MSP-1; merozoite surface protein; malaria; vaccine;
XX protein engineering; protein expression; codon usage;
XX transgenic animal; mutant.
XX OS Plasmodium falciparum.
XX Synthetic.
XX Key Location/Qualifiers
XX Peptide 1..15
XX /note="beta-casein signal peptide"
XX Peptide 371..376
XX /note="6xHis tag"
XX FT Misc-difference 197
XX /note="Asn in native MSP-1-42 (N181Q mutation)"
XX FT Misc-difference 278
XX /note="Asn in native MSP-1-42 (N262Q mutation)"
XX PN WO920774-A2.
XX 29-APR-1999.
XX PF 20-OCT-1998; 98WO-US22225.
XX PR 15-MAY-1998; 98US-0085649.
XX PR 20-OCT-1997; 97US-0062592.
XX PA (GENZ ) GENZYME TRANSGENICS CORP.
XX PI Chen LH, Meade H;
XX DR WPI; 1999-288313/24.
XX DR P-PSDB; AAX56008.
XX PT Modified malarial protein for use in anti-malarial vaccines
XX PS Example; Fig 11; 35pp; English.
XX CC The present sequence represents a 42 kDa C-terminal portion of
XX CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
XX CC important target for the development of a vaccine against
XX CC Plasmodium falciparum. The sequence has been modified to include
XX CC an N-terminal beta-casein signal peptide and an N-terminal 6xHis
XX CC tag. In addition, N181Q and N262Q mutations have been introduced to
XX CC eliminate 2 N-glycosylation sites. These modifications allow the
XX CC MSP-1-42 protein to be expressed in the milk of transgenic mice.
SQ Sequence 376 AA:

Query Match 93.0%; Score 1919.5; DB 20; Length 376;
Best Local Similarity 97.3%; Pred. No. 5.5e-127;
Matches 364; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

OY 12 VVYVATL-CTAGAAVPSYIDNLSKIEVEYLYKPLAGYRSIKKOLENNVMTFN 70
DB 3 VVIACTVLAATAVPSYIDNLSKIEVEYLYKPLAGYRSIKKOLENNVMTFN 62
OY 71 VVDIINSRKNKRNKFNVEDLIPKDLTSSNYYVKKPKLNKRKDKPLSSYVYKD 130
DB 63 VVDIINSRKNKRNKFNVEDLIPKDLTSSNYYVKKPKLNKRKDKPLSSYVYKD 122
OY 131 SIDTDFNANDVLYGYYKIISEKYSKSDLSIKRYINDKOGENEKYLPLNNIEFLK 190
DB 123 SIDTDFNANDVLYGYYKIISEKYSKSDLSIKRYINDKOGENEKYLPLNNIEFLK 182
OY 191 KIDLFYIHLEAKVLYNTYKSNVEYKIKELNLTKTODLADPKKNNNVGADLST 250
DB 191 KIDLFYIHLEAKVLYNTYKSNVEYKIKELNLTKTODLADPKKNNNVGADLST 250

```

```

DB 183 KIDLFYIHLEAKVLYNTYKSNVEYKIKELNLTKTODLADPKKNNNVGADLST 242
OY 251 HNNLTKRPLSTGVAFENLAKTVLSNLDGNLQMLNISOHCYKCKPONSQCFRHLDR 310
DB 243 HNNLTKRPLSTGVAFENLAKTVLSNLDGNLQMLNISOHCYKCKPONSQCFRHLDR 302
OY 311 EECCKLLNRYKQEGDKCVENPNPTCNENNGCCADAKCTEDSGSNKKITTCBCTK 370
DB 303 EECCKLLNRYKQEGDKCVENPNPTCNENNGCCADAKCTEDSGSNKKITTCBCTK 362
OY 371 LFDGIFCSHHHHH 384
DB 363 LFDGIFCSHHHHH 376

RESULT 4
AAY05834
ID AAY05834 standard; Protein; 376 AA.
XX AC AAY05834;
XX 02-AUG-1999 (first entry)
XX DE Modified merozoite surface protein MSP-1-42.
XX KW MSP-1; merozoite surface protein; malaria; vaccine;
XX protein engineering; protein expression; codon usage;
XX transgenic animal; mutant.
XX OS Plasmodium falciparum.
XX Synthetic.
XX Key Location/Qualifiers
XX Peptide 1..15
XX /note="beta-casein signal peptide"
XX Peptide 371..376
XX /note="6xHis tag"
XX FT Misc-difference 197
XX /note="Asn in native MSP-1-42 (N181Q mutation)"
XX FT Misc-difference 278
XX /note="Asn in native MSP-1-42 (N262Q mutation)"
XX PN WO920766-A2.
XX 29-APR-1999.
XX PF 20-OCT-1998; 98WO-US22225.
XX PR 15-MAY-1998; 98US-0085649.
XX PR 20-OCT-1997; 97US-0062592.
XX PA (GENZ ) GENZYME TRANSGENICS CORP.
XX PI Chen LH, Meade H;
XX DR WPI; 1999-302742/25.
XX DR N-PSDB; AAX25593.
XX PT New modified recombinant nucleic acid sequences useful for producing
XX PT malarial DNA vaccine
XX PS Example; Fig 11; 43pp; English.
XX CC The present sequence represents a 42 kDa C-terminal portion of
XX CC malaria merozoite surface protein MSP-1 (i.e. MSP-1-42), an
XX CC important target for the development of a vaccine against
XX CC Plasmodium falciparum. The sequence has been modified to include
XX CC an N-terminal beta-casein signal peptide and an N-terminal 6xHis
XX CC tag. In addition, N181Q and N262Q mutations have been introduced to
XX CC eliminate 2 N-glycosylation sites. These modifications allow the
XX CC MSP-1-42 protein to be expressed in the milk of transgenic mice.
XX CC The invention generally relates to modified recombinant nucleic
XX CC acid sequences and methods for increasing the mRNA levels and

```

protein expression of proteins that are difficult to express in cell culture systems, mammalian cell culture systems or in transgenic animals. The preferred difficult protein candidates for expression are those derived from lower organisms such as parasites, bacteria and viruses that have DNA coding sequences of high AT content or which have mRNA instability motifs or rare codons relative to the recombinant expression system to be used.

Sequence 376 AA;

Query Match 93.0%; Score 1919.5; DB 20; Length 376;
Best Local Similarity 97.3%; Pred. No. 5.5e-127;
Matches 364; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 12 VLVTATL-CTAGAATPSVIDNIIISKIENEYEVLYKPLAGVYSIKKOLENNVMTFNVN 70
D3 3 VLIACVLALAAVTPSVIDNIIISKIENEYEVLYKPLAGVYSIKKOLENNVMTFNVN 62
QY 71 VKDILNSRFRKREKRVNLESDLIPYKDLSSNTVYKDPYKFLNKKRDKFLSSYNTIKD 130
D3 63 VKDILNSRFRKREKRVNLESDLIPYKDLSSNTVYKDPYKFLNKKRDKFLSSYNTIKD 122
QY 131 SIDDIPFANDVLYGYKILSEKYSDDLIRKYYINDKGENEKYLPFLNTELYKYVNH 190
D3 123 SIDDIPFANDVLYGYKILSEKYSDDLIRKYYINDKGENEKYLPFLNTELYKYVND 182
QY 191 KIDFVHLEAKVLYTEKSNVEVKIKELNYLKTIDKLADEFKNNNFVGIADLSTDYN 250
D3 183 KIDFVHLEAKVLYTEKSNVEVKIKELNYLKTIDKLADEFKNNNFVGIADLSTDYN 242
QY 251 HNNILTRFLSTGMFENIAKTVLSNLDGNLQGMNLISQHCYVKQCPQNSGCFRHIDER 310
D3 243 HNNILTRFLSTGMFENIAKTVLSNLDGNLQGMNLISQHCYVKQCPQNSGCFRHIDER 302
QY 311 EECCLNTYKQEGKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTKPSYP 370
D3 303 EECCLNTYKQEGKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTKPSYP 362
QY 371 LFDGIFCSHHHHH 384
D3 363 LFDGIFCSHHHHH 376

RESULT 5

AAWS4145
ID AAWS4145 standard; Protein; 1639 AA.

AAWS4145;

DT 23-SEP-1998 (first entry)

DE P. falciparum synthetic gp190 protein.

KW gp190; malaria; MSP-1; merozoite surface protein; stability; vaccine; monoclonal antibody; passive immunisation; parasite.

OS Plasmodium falciparum.
Synthetic.

PN W09814583-A2.

PD 09-APR-1998.

PE 02-OCT-1997; 97WO-EP05441.

PR 02-OCT-1996; 96DE-4040817.

PA (BUJA/) BUJARD H.

PI Bujard H, Pan W, Tolle R;

PR MPI; 1998-240088/21.

DR N-PSDB; AAV21451, AAV35363.

XX Recombinant production of complete gp190/MSP-1 plasmodium surface protein - useful in anti-malaria vaccines, also stabilising genes by reducing their AT content

PS Example 1; Fig 3c; 48pp; German.

CC This sequence represents a modified plasmodium falciparum gp190/MSP-1 (merozoite surface) protein. The gene encoding this protein has been stabilised by reducing the AT content of the nucleotide sequence. Such a protein is useful in vaccines against malaria or for producing monoclonal antibodies (for passive immunisation). The complete gp190 protein can now be produced outside the parasite and has, at least over extended regions, the native pattern of folding. Larger amounts of the protein can be produced recombinantly than would be possible using the parasites as source.

Sequence 1639 AA;

Query Match 91.1%; Score 1879; DB 19; Length 1639;
Best Local Similarity 99.4%; Pred. No. 2.5e-123;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSVIDNIIISKIENEYEVLYKPLAGVYSIKKOLENNVMTFNVNKKDILNSRFRK 81
D3 1262 GAATPSVIDNIIISKIENEYEVLYKPLAGVYSIKKOLENNVMTFNVNKKDILNSRFRK 1321
QY 82 RENEKRVNLESDLIPYKDLSSNTVYKDPYKFLNKKRDKFLSSYNTIKDSIDDPINFAND 141
D3 1322 RENEKRVNLESDLIPYKDLSSNTVYKDPYKFLNKKRDKFLSSYNTIKDSIDDPINFAND 1381
QY 142 VLGYYKILSEKYSDDLIRKYYINDKGENEKYLPFLNTELYKYVNHKIDLFVHLEA 201
D3 1382 VLGYYKILSEKYSDDLIRKYYINDKGENEKYLPFLNTELYKYVNDKIDLFVHLEA 1441
QY 202 KVLNTEKSNVEVKIKELNYLKTIDKLADEFKNNNFVGIADLSTDYNNHNLTRFLST 261
D3 1442 KVLNTEKSNVEVKIKELNYLKTIDKLADEFKNNNFVGIADLSTDYNNHNLTRFLST 1501
QY 262 GMVFENIAKTVLSNLDGNLQGMNLISQHCYVKQCPQNSGCFRHIDERRECKCLLYKQ 321
D3 1502 GMVFENIAKTVLSNLDGNLQGMNLISQHCYVKQCPQNSGCFRHIDERRECKCLLYKQ 1561
QY 322 EGDKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTKPSYPLFDGIFCS 378
D3 1562 EGDKCVENPPTCNENNGGCDADAKCTEEDSGSNGKRTCECTKPSYPLFDGIFCS 1618

RESULT 6

AAI09372
ID AAI09372 standard; Protein; 355 AA.

AAI09372;

DT 31-AUG-1999 (first entry)

DE Merozoite surface protein MSP-1-42.

KW MSP-1; merozoite surface protein; malaria; vaccine; protein engineering; protein expression; codon usage; transgenic animal.

OS Plasmodium falciparum.

PN W09920774-A2.

PD 29-APR-1999.

PE 20-OCT-1998; 98WO-US22226.

PR 15-MAY-1998; 98US-0085649.

DR 20-OCT-1997; 97US-0062592.

PA (GEN2) GENZYME TRANSGENICS CORP.
 XX
 PI Chen LH, Meade H;
 XX
 DR WPI: 1999-288313/24.
 DR P-PSDB; AAX56008.
 XX
 PT Modified malarial protein for use in anti-malarial vaccines
 PS
 XX Example; Fig 1; 35pp; English.
 XX
 CC The present sequence represents a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (1.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42
 CC has been modified (see AAX56008) compared to the native sequence (see
 CC AAX56009) such that 306 nucleotide positions have been replaced to
 CC lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA
 CC instability motifs while maintaining the same protein amino acid
 CC sequence. These alterations allow MSP-1-42 to be expressed in
 CC mammalian cell culture and in transgenic mice. Native MSP-1-12
 CC is known to be difficult to express in cell culture systems,
 CC mammalian cell culture systems or in transgenic animals. The
 CC invention allows expression of MSP-1 protein in the milk of
 CC transgenic animals, and also provides a DNA vaccine comprising a
 CC vector containing the altered MSP-1-42 sequence.
 CC
 XX
 SQ Sequence 355 AA;
 Query Match 90.8%; Score 1874; DB 20; Length 355;
 Best Local Similarity 99.7%; Pred. No. 8.1e-124;
 Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 24 AATPSYIDNLSIKENEYEVLYKPLAGYRSLSKRLKLENNVMTFNNVNDILNSRPNKRE 83
 DB 1 AATPSYIDNLSIKENEYEVLYKPLAGYRSLSKRLKLENNVMTFNNVNDILNSRPNKRE 60
 OY 84 NFRNVLESDLIPKDLTSSNNYVKKDPYKFLNKKRDKFLSSYNYIKDSIDTDFINFANDYL 143
 DB 61 NFRNVLESDLIPKDLTSSNNYVKKDPYKFLNKKRDKFLSSYNYIKDSIDTDFINFANDYL 120
 OY 144 GYKILSEKYSKSLDSIKRYINDKOGENKRYLPFLNNITETVYVNHKIDLPYIHLEAVY 203
 DB 121 GYKILSEKYSKSLDSIKRYINDKOGENKRYLPFLNNITETVYVNHKIDLPYIHLEAVY 180
 OY 204 LNTTYEKSNEVEYKIKELANTLKTIDKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 263
 DB 181 LNTTYEKSNEVEYKIKELANTLKTIDKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 240
 OY 264 VRENIAKTVLSNLDGNLQGMNLISQHCVKKQCPONSGCFRHLDERECCKILANKQSG 323
 DB 241 VRENIAKTVLSNLDGNLQGMNLISQHCVKKQCPONSGCFRHLDERECCKILANKQSG 300
 OY 324 DKCVENPNTCENNGCGADAKCTBEDSGSNKRTTCTCTKPSDPLFDGIFCS 378
 DB 301 DKCVENPNTCENNGCGADAKCTBEDSGSNKRTTCTCTKPSDPLFDGIFCS 355
 RESULT 7
 AAY05832
 ID AAY05832 standard; Protein: 355 AA.
 XX
 AC AAY05832;
 XX
 DT 02-AUG-1999 (first entry)
 XX
 DE Merozoite surface protein MSP-1-42.
 XX
 KW MSP-1; merozoite surface protein; malaria; vaccine;
 KW protein engineering; protein expression; codon usage;
 KW transgenic animal.
 XX
 OS Plasmodium falciparum.

XX
 PN WO9920766-A2.
 XX
 PD 29-APR-1999.
 XX
 PF 20-OCT-1998; 98WO-US22225.
 XX
 PR 15-MAY-1998; 98US-0085649.
 PR 20-OCT-1997; 97US-0062592.
 XX
 PA (GEN2) GENZYME TRANSGENICS CORP.
 XX
 PI Chen LH, Meade H;
 XX
 DR WPI: 1999-302742/25.
 DR N-PSDB; AAX25586.
 XX
 PT New modified recombinant nucleic acid sequences useful for producing
 PT malarial DNA vaccine
 PS
 XX
 XX Disclosure; Fig 1; 43pp; English.
 CC The present sequence represents a 42 kDa C-terminal portion of
 CC malaria merozoite surface protein MSP-1 (1.e. MSP-1-42), an
 CC important target for the development of a vaccine against
 CC Plasmodium falciparum. The nucleic acid sequence encoding MSP-1-42
 CC has been modified (see AAX25586) compared to the native sequence (see
 CC AAX25587) such that 306 nucleotide positions have been replaced to
 CC lower the AT content (from 76 to 49.7%) and to eliminate 10 mRNA
 CC instability motifs while maintaining the same protein amino acid
 CC sequence. These alterations allow MSP-1-42 to be expressed in
 CC mammalian cell culture and in transgenic mice. The invention
 CC provides modified recombinant nucleic acid sequences and methods for
 CC increasing the mRNA levels and protein expression of proteins that
 CC are difficult to express in cell culture systems, mammalian cell
 CC culture systems or in transgenic animals. The preferred difficult
 CC protein candidates for expression are those derived from lower
 CC organisms such as parasites, bacteria and viruses that have DNA
 CC coding sequences of high AT content or which have mRNA instability
 CC motifs or rare codons relative to the recombinant expression system
 CC to be used. The invention allows expression of MSP-1 protein in
 CC the milk of transgenic animals, and also provides a DNA vaccine
 CC comprising a vector containing the altered MSP-1-42 sequence.
 CC
 XX
 SQ Sequence 355 AA;
 Query Match 90.8%; Score 1874; DB 20; Length 355;
 Best Local Similarity 99.7%; Pred. No. 8.1e-124;
 Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 24 AATPSYIDNLSIKENEYEVLYKPLAGYRSLSKRLKLENNVMTFNNVNDILNSRPNKRE 83
 DB 1 AATPSYIDNLSIKENEYEVLYKPLAGYRSLSKRLKLENNVMTFNNVNDILNSRPNKRE 60
 OY 84 NFRNVLESDLIPKDLTSSNNYVKKDPYKFLNKKRDKFLSSYNYIKDSIDTDFINFANDYL 143
 DB 61 NFRNVLESDLIPKDLTSSNNYVKKDPYKFLNKKRDKFLSSYNYIKDSIDTDFINFANDYL 120
 OY 144 GYKILSEKYSKSLDSIKRYINDKOGENKRYLPFLNNITETVYVNHKIDLPYIHLEAVY 203
 DB 121 GYKILSEKYSKSLDSIKRYINDKOGENKRYLPFLNNITETVYVNHKIDLPYIHLEAVY 180
 OY 204 LNTTYEKSNEVEYKIKELANTLKTIDKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 263
 DB 181 LNTTYEKSNEVEYKIKELANTLKTIDKLADFKKNNNFVGADISTDYNHNNLLTKFLSTGM 240
 OY 264 VRENIAKTVLSNLDGNLQGMNLISQHCVKKQCPONSGCFRHLDERECCKILANKQSG 323
 DB 241 VRENIAKTVLSNLDGNLQGMNLISQHCVKKQCPONSGCFRHLDERECCKILANKQSG 300
 OY 324 DKCVENPNTCENNGCGADAKCTBEDSGSNKRTTCTCTKPSDPLFDGIFCS 378
 DB 301 DKCVENPNTCENNGCGADAKCTBEDSGSNKRTTCTCTKPSDPLFDGIFCS 355

```

RESULT 8
AAP50777
ID AAP50777 standard; Protein: 1654 AA.
xx AAP50777;
xx 30-SEP-1991 (first entry)
xx Sequence of the P195 protein of Plasmodium falciparum.
xx
xx Malaria vaccine; epitope; antigen; immunogen.
xx
xx Plasmodium falciparum.
xx
xx EP154454-A.
xx
xx 11-SEP-1985.
xx
xx 21-FEB-1985; 85EP-0301173.
xx
xx 26-SEP-1984; 84GB-0024340.
xx 22-FEB-1984; 84GB-0004692.
xx 21-FEB-1985; 85GB-0004429.
xx
xx (WELL ) WELLCOME FOUNDATION LTD.
xx
xx Holder A, Sandhu J, Odink K, Lockyer M, Riveros-Moreno V;
xx WPI; 1985-224845/37.
xx N-PSDB; AAN50530.
xx
xx Cloned DNA sequence encoding plasmodium falciparum protein -
xx useful for expressing the protein for use in vaccines against
xx malaria
xx
xx
xx Claim 6; Fig 1; 51pp; English.
xx
xx The sequence encoding the P195 protein of Plasmodium falciparum
xx (AAN50530) and a peptide comprising at least one of its epitopes
xx (see AAP50777) are claimed. Also claimed is a vaccine for inducing
xx immunity to malaria comprising the novel peptide or P195 or a
xx peptide comprising at least one epitope when derived from the new
xx DNA sequence, together with a carrier.
xx
xx Sequence 1654 AA:
SO
Query Match 81.7%; Score 1684.5; DB 6; Length 1654;
Best Local Similarity 91.3%; Pred. No. 1.2e-109;
Matches 327; Conservative 6; Mismatches 24; Indels 1; Gaps 1;
QY 22 GAATPSPYINDILSKIENEYEVLYLKLPLAGYRSLSKQLENNVATFNVNVADILNSRPNK 81
DB 1276 GEAVTSTVIDNILSKIENEYEVLYLKLPLAGYRSLSKQLENNVATFNVNVADILNSRPNK 1335
QY 82 RERNFNVLSEDLIPEYKDLTSSNYVVKDPYKFLNKKRDKRFSSNYIKDSIDTDINFAND 141
DB 1336 RERNFNVLSEDLIPEYKDLTSSNYVVKDPYKFLNKKRDKRFSSNYIKDSIDTDINFAND 1395
QY 142 VLGYIKILSEKYSKSDLSIKRYINDKQGENEKYLPFLNNITETLYKYVNHKIDLFVYIHLEA 201
DB 1396 VLGYIKILSEKYSKSDLSIKRYINDKQGENEKYLPFLNNITETLYKYVNDKIDLFVYIHLEA 1455
QY KLVNATYKSVNVEYKIKELANTLKTIDOKLADFKKNNNVGVGADLSTQVNNHNLTKRFLST 261
DB 1456 KLVNATYKSVNVEYKIKELANTLKTIDOKLADFKKNNNVGVGADLSTQVNNHNLTKRFLST 1515
QY 262 GAVENELAKTYVLSNLDGNT-QGMNLNTSOHQCVKRCOPNSGCRPHLDERECKCLINLYK 320
DB 1516 GAVENELAKTYVLSNLDGNTKRLARVYKHHFTTPRRKTKMTIQOSSGCRPHLDERECKCLINLYK 1575
QY 321 QEGKRCVYENPPTCGENNGGCDADAKCTEEDSGSNGKKITYCECTKPPSDYPLFDGIFCS 378

```

Db	15716	QEGSKCVENSNPTCHENNCGGDADAKCTEEDSGSNGKRTTCECTCFPDGPIPLSMVIFCS	1633
Db	15716	QEGSKCVENSNPTCHENNCGGDADAKCTEEDSGSNGKRTTCECTCFPDGPIPLSMVIFCS	1633
XX	RESULT 9		
XX	AAB37610		
XX	AAB37610 standard; Protein: 262 AA.		
XX	AAB37610;		
XX	27-FEB-2001 (first entry)		
XX	Merozoite surface protein-133.		
XX	Merozoite surface protein; protazoacide; vaccine; malaria.		
XX	Plasmodium falciparum.		
XX	W0200063245-A2.		
XX	26-OCT-2000.		
XX	20-APR-2000; 2000WO-GB01558.		
XX	20-APR-1999; 99GB-0009072.		
XX	13-MAY-1999; 99US-0311817.		
XX	25-MAY-1999; 99CA-2271451.		
XX	(MEDI-) MEDICAL RES COUNCIL.		
XX	Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthaijibull C;		
XX	WPI: 2001-015762/02.		
XX	N-PSDB; AAC68978.		
XX	Novel variants of the C-terminal fragment of Plasmodium merozoite		
XX	surface protein-1, useful as vaccines for treating or preventing		
XX	malaria		
XX	Example 5; Fig 15; 126pp; English.		
XX	The present invention relates to non-natural variants of a C-terminal		
XX	fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The		
XX	non-natural variants have reduced affinity for at least 1 antibody		
XX	capable of blocking a second antibody that inhibits the proteolytic		
XX	cleavage of Plasmodium MSP-1_4_2, and has the same affinity for at least		
XX	one third antibody that inhibits the proteolytic cleavage of Plasmodium		
XX	MSP-1_4_2, compared to natural MSP-1_1_9. The non-natural variants of the		
XX	present invention are useful for immunising a mammal against malaria, and		
XX	can be used to treat malaria. The present sequence is MSP-133 protein.		
XX	Sequence 262 AA;		
XX	Query Match 64.3%; Score 1326; DB 22; Length 262;		
XX	Best Local Similarity 99.2%; Pred. No. 1.9e-85;		
XX	Matches 260; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
XX	24 AAVPSVIDNLSIKTENEEVLYLKPIAGVYSIKRKOLENNMTFVNVNKDILNSFNRE 83		
XX	1 AAVTSVIDNLSIKTENEEVLYLKPIAGVYSIKRKOLENNMTFVNVNKDILNSFNRE 60		
XX	84 NFKNVLSDDLIPYKDLTSSNVVADPYKFLNKKERDKFLSSYNIKSIDIDINFAVDYL 143		
XX	61 NFKNVLSDDLIPYKDLTSSNVVADPYKFLNKKERDKFLSSYNIKSIDIDINFAVDYL 120		
XX	144 GYKIIILSEKYSKSLDLSIKKIYINDKOGENEKTLPELNNIETLYKTVNKKIDLPVHLKAV 203		
XX	121 GYKIIILSEKYSKSLDLSIKKIYINDKOGENEKTLPELNNIETLYKTVNKKIDLPVHLKAV 180		
XX	204 LNVYKESNVVYKIKELNYLTKTIDDKLADPFKNNNPGVIALSDYNNNNLITKFLSTGCM 263		
XX	181 LNVYKESNVVYKIKELNYLTKTIDDKLADPFKNNNPGVIALSDYNNNNLITKFLSTGCM 240		

OY 264 VFEINAKTVLSNLDGNLQGM 285
 DB 241 VFEINAKTVLSNLDGNLQGM 262

RESULT 10

AAB83926
 ID AAB83926 standard; Protein; 375 AA.

AC AAB83926;

DT 23-JUL-2001 (first entry)

DE A major merozoite surface protein-1 fragment of 42kDa.

KW Major merozoite surface protein-1; MSP-142; melltlin signal peptide; malaria vaccine.

OS Plasmodium falciparum.

PN WO200134188-A1.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000MO-US31064.

PR 12-NOV-1999; 99US-0165178.

PR 01-DEC-1999; 99US-0168327.

PR 22-AUG-2000; 2000US-0226861.

PA (UYHA-) UNIV HAWAII.

PA (UYCH-) UNIV CHINESE HONG KONG.

PA (QUEB-) QUEEN EMMA FOUND.

PI Hui GSN, Lap-yin P, Ho WK;

DR WPI; 2001-335879/35.

DR N-PSDB; AAF89840.

XX Producing malaria vaccine, useful for treatment or prevention of all

XX PT forms of malaria in humans, by expressing immunogenic merozoite protein

XX PT fragment in a baculovirus system

XX PS Example 3; Page 87-88; 95pp; English.

XX The present sequence represents a major merozoite surface protein-1

XX CC C-terminal fragment of 42kDa (MSP-142). This fragment is linked to a

XX CC melltlin signal peptide, and then expressed in a in a

XX CC silworm/baculovirus system. The protein is used to prepare a

XX CC malaria vaccine, which is used to treat or prevent malaria, caused by

XX CC any of the four species of Plasmodium that infect humans.

XX SQ Sequence 375 AA;

OY Query Match

DB Best Local Similarity 56.3%; Score 1161.5; DB 22; Length 375;

DB Matches 218; Conservative 54; Mismatches 76; Indels 21; Gaps 2;

OY 30 INIISIKINEYEVLYIKPLAGVYSLKQLENNVTFVNVKDLNSFRNENKYL 89

DB 7 MNINISGFENEYDYLKPLAGVYSLKQLENNVTFVNVKDLNSFRNENKYL 66

OY 90 ESDLIPYKDLSSNVVYKPLNKKRDLSSNYTKSDIDPINDVAGYKIL 149

DB 67 ESDLIPYKDLSSNVVYKPLNKKRDLSSNYTKSDIDPINDVAGYKIL 126

OY 127 LAKYDDLESIKKVKKEKEKPPSPPTPPSPAKTDEQKESKELPFLTNIETLYNNLY 186

DB 127 LAKYDDLESIKKVKKEKEKPPSPPTPPSPAKTDEQKESKELPFLTNIETLYNNLY 186

OY 190 HKIDFVHLLEAKVNLVYTKSNVYKIKELVYTKTIOKDLADFKNNNPFVGIADISTYX 249

DB 187 NKIDYLLNKKAKINDCNVEKDEAHVYKTKLSLDAIDKIDLFPKPYDFEAIKILIND 246

OY 250 NNNLLTKFELSTGMEFENIAKTVLSNLDGNLQGMNLSQHOCVCKOCPPONGCFRRHDE 309

DB 247 TKKDMGKILSTGLV-QNFPNIIISKILIGKFDQDMNLSQHOCVCKOCPPONGCFRRHDE 305

OY 310 RECKCLLNKYKQEGDCEVNPPTCNENNGGCDADAKCTEEDSGSNKRTCTCPDSY 369

DB 306 RECKCLLNKYKQEGDCEVNPPTCNENNGGCDADAKCTEEDSGSNKRTCTCPDSY 365

OY 370 PLEDFGFC 378

DB 366 PLEDFGFC 374

RESULT 11

AAB37609
 ID AAB37609 standard; Protein; 108 AA.

AC AAB37609;

DT 27-FEB-2001 (first entry)

DE Merozoite surface protein-119.

KW Merozoite surface protein; protazoacide; vaccine; malaria.

OS Plasmodium falciparum.

PN WO200063245-A2.

PD 26-OCT-2000.

PF 20-APR-2000; 2000MO-GB01558.

PR 20-APR-1999; 99GB-0009072.

PR 13-MAY-1999; 99US-0311817.

PR 25-MAY-1999; 99CA-2271451.

PA (MED-) MEDICAL RES COUNCIL.

PI Holder A, Bidsall B, Feeney J, Morgan W, Syed S, Uthaiybull C;

DR WPI; 2001-015762/02.

DR N-PSDB; AAC68977.

XX Novel variants of the C-terminal fragment of Plasmodium merozoite

XX PT surface protein-1, useful as vaccines for treating or preventing

XX PT malaria

XX PS Example 5; Fig 15; 126pp; English.

XX The present invention relates to non-natural variants of a C-terminal

XX CC fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The

XX CC non-natural variants have reduced affinity for at least 1 antibody

XX CC capable of blocking a second antibody that inhibits the proteolytic

XX CC cleavage of Plasmodium MSP-1-4.2, and has the same affinity for at least

XX CC one third antibody that inhibits the proteolytic cleavage of Plasmodium

XX CC MSP-1-4.2, compared to natural MSP-1-1.9. The non-natural variants of the

XX CC present invention are useful for immunising a mammal against malaria, and

XX CC can be used to treat malaria. The present sequence is MSP-119 protein.

XX SQ Sequence 108 AA;

OY Query Match

DB Best Local Similarity 26.5%; Score 546; DB 22; Length 108;

DB Matches 93; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 280 NLOGLNLSQHOCVCKOCPPONGCFRRHDERECKCLLNKYKQEGDCEVNPPTCNENNG 339

DB 7 HIEGRNINLQHCQVCKOCPPONGCFRRHDERECKCLLNKYKQEGDCEVNPPTCNENNG 66

OY 340 GCDADAKCTEEDSGSNKRTCTCPDSYPLEDFGFC 378

DB 340 GCDADAKCTEEDSGSNKRTCTCPDSYPLEDFGFC 378

D6 67 GCDADAKCTEEDSGSGNGKKTTCCTKPDSTPLFDGIFCS 105

RESULT 12

AAAB37608
ID AAB37608 standard; protein; 96 AA.

AAAB37608;

27-FEB-2001 (first entry)

XX Merozoite surface protein-1.

XX Merozoite surface protein; protazoacide; vaccine; malaria.

OS Plasmodium falciparum.

XX WO200063245-A2.

26-OCT-2000.

20-APR-2000; 2000WO-GB01558.

20-APR-1999; 99GB-0009072.

13-MAY-1999; 98US-0311817.

25-MAY-1999; 99CA-2271451.

(MEDI-) MEDICAL RES COUNCIL.

Holder A, Birdsall B, Feeney J, Morgan W, Syed S, Uthathipibull C;

WPI; 2001-015762/02.

Novel variants of the C-terminal fragment of Plasmodium merozoite

surface protein-1, useful as vaccines for treating or preventing

malaria.

Example 2; Page 48; 126pp; English.

The present invention relates to non-natural variants of a C-terminal

fragment of a Plasmodium merozoite surface protein-1 (MSP-1). The

non-natural variants have reduced affinity for at least 1 antibody

capable of blocking a second antibody that inhibits the proteolytic

cleavage of Plasmodium MSP-1₄₋₂, and has the same affinity for at least

one third antibody that inhibits the proteolytic cleavage of Plasmodium

MSP-1₄₋₂, compared to natural MSP-1₁₋₉. The present sequence is the

wild-type MSP-1 protein. This sequence was used to generate the variants

of the present invention. The non-natural variants of the present

invention are useful for immunising a mammal against malaria, and can be

used to treat malaria.

Sequence 96 AA;

Query Match 26.3%; Score 542; DB 22; Length 96;

Best Local Similarity 100.0%; Pred. No. 6.9e-31;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

286 NISQHCYKQCCPQNSGCFRHLDERECCKLNTYKQGGKCVNPPTCENNNGGCDADA 345

1 NISQHCYKQCCPQNSGCFRHLDERECCKLNTYKQGGKCVNPPTCENNNGGCDADA 60

346 KCTEEDSGSGNGKKTTCCTKPDSTPLFDGIFCS 378

61 KCTEEDSGSGNGKKTTCCTKPDSTPLFDGIFCS 93

RESULT 13

AAAB36103

ID AAB36103 standard; protein; 116 AA.

AAAB36103;

25-MAR-1998 (first entry)

XX pFMSPI(p19)A protein sequence.

XX Plasmodium vivax; merozoite surface protein; MSP1; p19;

XX Plasmodium falciparum; malaria; vaccine; immunity; epitope.

XX Plasmodium falciparum.

OS Synthetic.

XX Key

FT Region 1.95 Location/Qualifiers

FT Region /note="amino acids derived from P. falciparum MSP1 p19

FT Region 96.116 fragment"

FT Region /note="glycosylphosphatidylinositol anchoring sequence"

XX WO9730158-A2.

21-AUG-1997.

14-FEB-1997; 97WO-FR00290.

14-FEB-1996; 96FR-0001822.

(INSP) INST PASTEUR.

(UNY) UNIV NEW YORK STATE.

Barnwell JM, Longacre-Andre S, Mendis K, Nato F;

WPI; 1997-425033/39.

Recombinant protein containing the merozoite surface protein-1 p19

fragment - useful in anti-malarial vaccines, diagnosis and protein

purification

Disclosure; Fig 1B; 85pp; French.

This is the amino acid sequence of a recombinant protein comprising

amino acids 1613-1705 of the Plasmodium falciparum merozoite surface

protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a

glycosylphosphatidylinositol membrane anchoring sequence. p19 is the

C-terminal fragment of the 42 kD MSP1 from Plasmodium species.

The recombinant protein can be used for the production of anti-malarial

vaccines, where the p19 fragment provides a high level of protective

immunity since it includes epitopes not presented in the p42 fragment.

Sequence 116 AA;

Query Match 26.1%; Score 539; DB 18; Length 116;

Best Local Similarity 98.9%; Pred. No. 1.4e-30;

Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

286 NISQHCYKQCCPQNSGCFRHLDERECCKLNTYKQGGKCVNPPTCENNNGGCDADA 345

3 NISQHCYKQCCPQNSGCFRHLDERECCKLNTYKQGGKCVNPPTCENNNGGCDADA 62

346 KCTEEDSGSGNGKKTTCCTKPDSTPLFDGIFCS 378

63 KCTEEDSGSGNGKKTTCCTKPDSTPLFDGIFCS 95

RESULT 14

AAAB22592

ID AAB22592 standard; protein; 116 AA.

AAAB22592;

25-MAR-1998 (first entry)

pFMSPI(p19)A protein sequence.

Key	Location/Qualifiers
Key	1..95
Region	/note= "amino acids derived from P. falciparum MSP1 p19 fragment"
Region	96..116
Region	/note= "glycosylphosphatidylinositol anchoring sequence"
MO9730159-A2.	
21-AUG-1997.	
14-FEB-1997;	97WO-FR00291.
14-FEB-1996;	96FR-0001821.
(INSP) INST PASTEUR.	
(UYNV) UNIV NEW YORK STATE.	
Barnewell JW, Longacre-Andre S, Mendis K, Nato F, Roth C;	
WPI; 1997-425034/39.	
P-PSDB; AAM22592.	
Recombinant protein containing Plasmodium mezoizote surface protein-1 p42 fragment - useful in antimalarial vaccines, also new antibodies for diagnosis and protein purification	
Disclosure; Fig 1B; 85pp; French.	
This is the amino acid sequence of a recombinant protein comprising amino acids 1613-1705 of the Plasmodium falciparum mezoizote surface protein 1 (MSP1) 19 kD C-terminal fragment (p19), linked to a glycosylphosphatidylinositol membrane anchoring sequence. p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species. The recombinant protein can be used for the production of anti-malarial vaccines, where the p19 fragment provides a high level of protective immunity since it includes epitopes not presented in the p42 fragment.	
Sequence 116 AA;	
Query Match 26.1%; Score 539; DB 18; Length 116;	
Best Local Similarity 98.9%; Pred. No. 1.4e-30;	
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0	
286 NISOHQVKKQCPONSGCFRHLDERBECKCLINTYKQESDKCVENPPTCENNGCCADA 345	
3 NISOHCQKKCKPENSQCFRHLDERBECKCLINTYKQESDKCVENPPTCENNGCCADA 62	
346 KCTEDSGSGNGKKTICCTCKDYSVLPFGITCS 378	
63 KCTEDSGSGNGKKTICCTCKDYSVLPFGITCS 95	
RESULT 15	
AAM22593	
AAM22593 standard; Protein; 127 AA.	
AAM22593;	
25-MAR-1998 (first entry)	
PfMSP1(p19)S protein sequence.	
Chimeric; Plasmodium vivax; mezoizote surface protein; MSP1; p19;	
Plasmodium falciparum; malaria; vaccine; immunity; epitope.	

```

OS Chimeric Plasmodium vivax.
OS Chimeric - Plasmodium falciparum.
XX Key Location/Qualifiers
FH Peptide 1..19
FT /note= "signal peptide"
FT Protein 20..127
FT Region /note= "mature protein"
FT Region 1..32
FT /note= "derived from P. vivax MSP1"
FT Region 33..34
FT /note= "encoded by restriction enzyme sequence used to
FT create the chimeric sequence"
FT Region 35..127
FT /note= "derived from P. falciparum C-terminal p19
FT fragment of MSP1."
PN MO9730159-A2.
PD 21-AUG-1997.
XX
XX 14-FEB-1997; 97WO-FR00291.
PR 14-FEB-1996; 96FR-0001821.
XX
XX (INSP ) INST PASTEUR.
PA (UNYNY ) UNIV NEW YORK STATE.
XX
XX Barnwell JW, Longacre Andre S, Mendis K, Nato F,
PI Roth C;
PL
DR WP1: 1997-425034/39.
DR P-PSDB; AAW22592.
XX
XX Recombinant protein containing Plasmodium mezoite surface
PT protein-1 p42 fragment - useful in antimalarial vaccines, also new
PT antibodies for diagnosis and protein purification
XX
XX Disclosure; Fig 1C; 85pp; French.
XX
XX This is the amino acid sequence of a chimeric protein comprising amino
CC acids 1-32 of the Plasmodium vivax mezoite surface protein 1 (MSP1)
CC linked to amino acids Asn1613-Ser1705 which correspond to the 19 kD
CC C-terminal fragment (p19) of MSP-1 from Plasmodium falciparum.
CC p19 is the C-terminal fragment of the 42 kD MSP1 from Plasmodium species.
CC The recombinant protein can be used for the production of anti-malarial
CC vaccines, where the p19 fragment provides a high level of protective
CC immunity since it includes epitopes not presented in the p42 fragment.
XX
XX Sequence 127 AA:
SQ
Query Match 26.1%; Score 539; DB 18; length 127;
Best Local Similarity 98.9%; Pred. No. 1.6e-30;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 286 NISOHCKKRCPPONGSGCFRHLDEREBCKLINTYKQEDDKCVENPNCNENNGSCDDA 345
DB 35 NISOHCKKRCPPENGSGCFRLHDREBECKCLINTYKQEDDKCVENPNCNENNGSCDDA 94
OY 346 KCTEEDSGSNGRKITCECTKRPDSYPLFDGIFCS 378
DB 95 KCTEEDSGSNGRKITCECTKRPDSYPLFDGIFCS 127

```

THIS PAGE BLANK (00710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:03:10 ; Search time 27 Seconds
(without alignments)
418.459 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSWKCLPMAVLVATLCT.....KPDYPLFDGIFCSHHHHH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1719.5	83.3	377	4	US-08-195-705-5 Sequence 5, Appl1
2	1658.5	80.4	375	4	US-08-195-705-3 Sequence 3, Appl1
3	1191.5	57.8	394	4	US-08-195-705-2 Sequence 2, Appl1
4	1156.5	56.1	394	4	US-08-195-705-4 Sequence 4, Appl1
5	287	13.9	53	1	US-08-290-919-3 Sequence 3, Appl1
6	280	13.6	48	1	US-08-290-919-12 Sequence 12, Appl1
7	274	13.3	48	1	US-08-290-919-1 Sequence 1, Appl1
8	271	13.1	48	1	US-08-290-919-2 Sequence 2, Appl1
9	268	13.0	53	1	US-08-290-919-4 Sequence 4, Appl1
10	253.5	12.3	106	1	US-08-290-919-11 Sequence 11, Appl1
11	144	7.0	816	4	US-07-640-029-1 Sequence 1, Appl1
12	141.5	6.9	384	4	US-09-134-001C-4669 Sequence 4669, Ap
13	141.5	6.9	817	1	US-07-640-029-2 Sequence 2, Appl1
14	141.5	6.9	820	1	US-08-166-717D-6 Sequence 6, Appl1
15	141.5	6.9	822	1	US-07-997-133-1 Sequence 1, Appl1
16	141.5	6.9	822	1	US-07-921-807B-4 Sequence 4, Appl1
17	141.5	6.9	822	1	US-08-459-296-2 Sequence 2, Appl1
18	141.5	6.9	822	1	US-08-441-944A-4 Sequence 4, Appl1
19	141.5	6.9	822	2	US-08-451-822A-12 Sequence 12, Appl1
20	141.5	6.9	822	4	US-08-439-992A-2 Sequence 2, Appl1
21	141.5	6.9	822	1	US-08-323-430-12 Sequence 12, Appl1
22	138.5	6.7	820	1	US-07-921-807B-3 Sequence 3, Appl1
23	138.5	6.7	820	1	US-08-441-944A-3 Sequence 3, Appl1
24	138.5	6.7	820	4	US-08-439-992A-1 Sequence 1, Appl1
25	136	6.6	2391	4	US-08-446-855A-2 Sequence 2, Appl1
26	136	6.6	2391	4	US-09-150-741-2 Sequence 2, Appl1
27	132.5	6.4	455	2	US-08-392-625-21 Sequence 21, Appl1

28	132.5	6.4	455	2	US-08-466-961A-21 Sequence 21, Appl1
29	132.5	6.4	455	2	US-08-645-193B-23 Sequence 23, Appl1
30	131	6.3	520	4	US-09-752-165-1 Sequence 1, Appl1
31	129	6.3	302	1	US-07-640-029-6 Sequence 6, Appl1
32	129	6.3	302	1	US-07-921-807B-8 Sequence 8, Appl1
33	129	6.3	302	1	US-08-441-944A-8 Sequence 8, Appl1
34	129	6.3	302	4	US-08-439-992A-6 Sequence 6, Appl1
35	129	6.3	733	1	US-07-640-029-4 Sequence 4, Appl1
36	129	6.3	733	1	US-07-921-807B-6 Sequence 6, Appl1
37	129	6.3	733	4	US-08-441-944A-6 Sequence 6, Appl1
38	129	6.3	733	4	US-08-439-992A-4 Sequence 4, Appl1
39	128	6.2	731	1	US-07-921-807B-5 Sequence 5, Appl1
40	128	6.2	731	1	US-08-441-944A-5 Sequence 5, Appl1
41	128	6.2	731	4	US-08-439-992A-3 Sequence 3, Appl1
42	126	6.1	300	4	US-07-640-029-5 Sequence 5, Appl1
43	126	6.1	300	4	US-08-439-992A-5 Sequence 5, Appl1
44	126	6.1	729	1	US-07-640-029-3 Sequence 3, Appl1
45	124.5	6.0	1038	4	US-09-541-782-4 Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-08-195-705-5
; Sequence 5, Application US/08195705
; Patent No. 6420523
GENERAL INFORMATION:
APPLICANT: Chang, Sandra
APPLICANT: Hui, George
APPLICANT: Barr, Philip
APPLICANT: Gibson, Helen
TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
TITLE OF INVENTION: FALCIPARUM VACCINE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Davis Hoxie Faithfull Hapgood
STREET: 45 Rockefeller Pl.
CITY: New York
STATE: N.Y.
COUNTRY: USA
ZIP: 10111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,705
FILING DATE: 14-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jacobs, Seth H
REGISTRATION NUMBER: 32140
REFERENCE/DOCKET NUMBER: 11880A3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-757-2200
TELEFAX: 212-586-1461
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
STRAIN: WEL
US-08-195-705-5
Query Match 83.3%; Score 1719.5; DB 4; Length 377;

Best Local Similarity 93.08; Pred. No. 1,1e-121;
Matches 331; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

```

OY 24 AVPSVYDNLISLKIENEYEVLYLKLPLAGVYRSLSKQLENNVMTFNVVKDILNSRFKRE 83
DB 1 AVPSVYDNLISLKIENEYEVLYLKLPLAGVYRSLSKQLENNVMTFNVVKDILNSRFKRE 60
OY 84 NFKNVLESDILPYKDLTSSNVVADPYKFLNKKERKDFLSYNTIKSDIPTDINFANDVL 143
DB 61 NFKNVLESDILPYKDLTSSNVVADPYKFLNKKERKDFLSYNTIKSDIPTDINFANDVL 120
OY 144 GYKILSEKYSKSDLSIKKTIYNDKQGENEKLPLPLNNIETLYKTVNKKIDLFVYHLEAKV 203
DB 121 GYKILSEKYSKSDLSIKKTIYNDKQGENEKLPLPLNNIETLYKTVNKKIDLFVYHLEAKV 180
OY 204 LNTYKESNVEVKIKEINLYLKTIDOKLADFKKNNFVGIADLSTDYNNHNLTKFSLTGM 263
DB 181 LNTYKESNVEVKIKEINLYLKTIDOKLADFKKNNFVGIADLSTDYNNHNLTKFSLTGM 240
OY 264 VFNILATVLSNLDGML-OGMLNISOHCYKQCPONSGCFRHLDERECCILNTKOE 322
DB 241 VFNILATVLSNLDGMLARVYKFTTPMRKKTMIQSSGCFRHLDERECCILNTKOE 300
OY 323 GDCVKNPNPTCNENNGCCADACCTEDSGSNKKITTCCTKPDYPLFDGIFCS 378
DB 301 GDCVKNPNPTCNENNGCCADACCTEDSGSNKKITTCCTKPDYPLFDGIFCS 356

```

RESULT 2
US-08-195-705-3
Sequence 3, Application US/08195705
Patent No. 6420523

```

: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra
: APPLICANT: Hul, George
: APPLICANT: Barr, Philip
: APPLICANT: Gibson, Helen
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESSES:
: ADDRESSSEE: Davis Hoxie Faithfull Hapgood
: STREET: 45 Rockefeller Pl.
: CITY: New York
: STATE: N.Y.
: COUNTRY: USA
: ZIP: 10111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/195,705
: FILING DATE: 14-FEB-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jacobs, Seth H
: REGISTRATION NUMBER: 32140
: REFERENCE/DOCKET NUMBER: 11880A3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-757-2200
: TELEFAX: 212-586-1461
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 375 amino acids
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHEICAL: NO
: FRAGMENT TYPE: C-terminal
: ORIGINAL SOURCE:

```

ORGANISM: Plasmidium falciparum
STRAIN: K1
US-08-195-705-3

Query Match 80.48; Score 1658.5; DB 4; Length 375;
Best Local Similarity 90.78; Pred. No. 4.4e-117;
Matches 323; Conservative 7; Mismatches 23; Indels 3; Gaps 2;

```

OY 24 AVPSVYDNLISLKIENEYEVLYLKLPLAGVYRSLSKQLENNVMTFNVVKDILNSRFKRE 83
DB 1 AVPSVYDNLISLKIENEYEVLYLKLPLAGVYRSLSKQLENNVMTFNVVKDILNSRFKRE 60
OY 84 NFKNVLESDILPYKDLTSSNVVADPYKFLNKKERKDFLSYNTIKSDIPTDINFANDVL 143
DB 61 NFKNVLESDILPYKDLTSSNVVADPYKFLNKKERKDFLSYNTIKSDIPTDINFANDVL 120
OY 144 GYKILSEKYSKSDLSIKKTIYNDKQGENEKLPLPLNNIETLYKTVNKKIDLFVYHLEAKV 203
DB 121 GYKILSEKYSKSDLSIKKTIYNDKQGENEKLPLPLNNIETLYKTVNKKIDLFVYHLEAKV 180
OY 204 LNTYKESNVEVKIKEINLYLKTIDOKLADFKKNNFVGIADLSTDYNNHNLTKFSLTGM 263
DB 181 LNTYKESNVEVKIKEINLYLKTIDOKLADFKKNNFVGIADLSTDYNNHNLTKFSLTGM 238
OY 264 VFNILATVLSNLDGML-OGMLNISOHCYKQCPONSGCFRHLDERECCILNTKOE 322
DB 239 VFNILATVLSNLDGMLARVYKFTTPMRKKTMIQSSGCFRHLDERECCILNTKOE 298
OY 323 GDCVKNPNPTCNENNGCCADACCTEDSGSNKKITTCCTKPDYPLFDGIFCS 378
DB 299 GDCVKNPNPTCNENNGCCADACCTEDSGSNKKITTCCTKPDYPLFDGIFCS 354

```

RESULT 3
US-08-195-705-2

Sequence 2, Application US/08195705
Patent No. 6420523

```

: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra
: APPLICANT: Hul, George
: APPLICANT: Barr, Philip
: APPLICANT: Gibson, Helen
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESSES:
: ADDRESSSEE: Davis Hoxie Faithfull Hapgood
: STREET: 45 Rockefeller Pl.
: CITY: New York
: STATE: N.Y.
: COUNTRY: USA
: ZIP: 10111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/195,705
: FILING DATE: 14-FEB-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jacobs, Seth H
: REGISTRATION NUMBER: 32140
: REFERENCE/DOCKET NUMBER: 11880A3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-757-2200
: TELEFAX: 212-586-1461
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 394 amino acids
: TYPE: amino acid
: STRANDEDNESS: single

```

```

:      TOPOLOGY: linear
:      MOLECULE TYPE: peptide
:      HYPOTHEICAL: NO
:      ANTI-SENSE: NO
:      FRAGMENT TYPE: C-terminal
:      ORIGINAL SOURCE:
:      ORGANISM: Plasmodium falciparum
:      STRAIN: falciparum uganda palo alto (FUP)
US-08-195-705-2

Query Match      57.8%; Score 1191.5; DB 4; Length 394;
Best Local Similarity 60.2%; Pred. No. 5,6e-82;
Matches 222; Conservative 54; Mismatches 72; Indels 21; Gaps 2

OY      30 IDNLSKRENEVEYVLYKPLAGVYRSLSKROLENNYMTFVNVYKDLNLSRFRNKREKNYL 89
      DB      6 MDNLSGSENERDYVLYKPLAGVYRSLSKQLEKNTFTNLNNDLNSRLKRRKFTLV 65
OY      90 ESDLLPYKDLTSSNVYVDPYKFLNKKERDKFLSSYNYIKDSIDDINDPANDVLYGYKTL 149
      DB      66 ESDLLQGFHSSINEYIIEDSEFLNSEQNTLTKSYKYIKESVENDIKFAGGISYEVK 125
OY      150 SEKKSDLSIKKYI-----NDQGENEKLPLFNINLETLYKTYN 189
      DB      126 LAKTKDLESIKKVKIEKEKRPSSPPTPPSPATDQCKSKSLPLFTNLETLYNNLY 185
OY      190 HKIDFLVHLEKVLNTYYSKNVYKIKELNYLTIDKLDLDFKNNFVGIADLSTDY 249
      DB      186 NKIDYDILNKAINDCVENDEAHVKTITKLSDLAIDKIDOLFNNHDFEAIKTLIND 245
OY      250 NNNLLTFLSTGWFEMLAKTVLSNLDGNIQGMINTSIQHCYKCKQPSNGCFRHDE 309
      DB      246 TKRMDLGRILSTGLV-QNFPNTIISKLIEGKFDMDLINTSIQHCYKCKQPSNGCFRHDE 304
OY      310 REECKCLINTYKQEGKCYENRPTCNENNGCGDADAKCTEESGSGNGKRTTCEYRPSY 369
      DB      305 REECKCLINTYKQEGKCYENRPTCNENNGCGDADAKCTEESGSGNGKRTTCEYRPSY 364
OY      370 PLFDGIFCS 378
      DB      365 PLFDGIFCS 373

RESULT 4
US-08-195-705-4
: Sequence 4, Application US/08195705
: Patent No. 6420523
: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra
: APPLICANT: Hul, George
: APPLICANT: Barr, Philip
: APPLICANT: Gibson, Helen
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM
: TITLE OF INVENTION: FALCIPARUM VACCINE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Davis Hoxie Faithfull Hapgood
: STREET: 45 Rockefeller Pl.
: CITY: New York
: STATE: N.Y.
: COUNTRY: USA
: ZIP: 10111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/195,705
: FILING DATE: 14-FEB-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jacobs, Seth H

```

```

1  REGISTRATION NUMBER: 32140
2  REFERENCE/DOCKET NUMBER: 1188043
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: 212-757-2200
5  TELEFAX: 212-586-1461
6  INFORMATION FOR SEQ ID NO: 4:
7  SEQUENCE CHARACTERISTICS:
8  LENGTH: 394 amino acids
9  TYPE: amino acid
10 STRANDEDNESS: single
11 TOPOLOGY: linear
12 MOLECULE TYPE: peptide
13 HYPOTHEICAL: NO
14 FRAGMENT TYPE: C-terminal
15 ORIGINAL SOURCE:
16 ORGANISM: Plasmodium falciparum
17 STRAIN: MAD
18 US-08-195-705-4
19
20 Query Match 56.1%; Score 1156.5; DB 4; Length 394;
21 Best Local Similarity 59.0%; Pred. No. 2.4e-79;
22 Matches 219; Conservative 54; Mismatches 73; Indels 25; Gaps 3
23
24 QY 30 IDNLSKTIENYEVLVIKPLAGVYRSLSKKOLENNVTFENVYKDIILNSFRKNREPKNYL 89
25      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
26 6 MDNLSGEGNEYDYVILKPLAGVYRSLSKKOLEKNTITFENLNDILNSRLKRRKYFLDYL 65
27      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 QY 90 ESDLIPIYDILSSNVYVDPYKFLNKKERDKRDFLSSYNTIKDSIDIDIDINFANDVIGYKTL 149
29      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30 66 ESDLMQFHNHSSNETIIEDSFLLNSEOKNILLKSYKIKESVENDIDFAQGISYIEYV 125
31      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 QY 150 SEKYSKSDSIRKKY-----NDXGENSEKYLPIFLNIEETLYKTVN 189
33      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 Db 126 LAKFKDLESIKKVIKEKEKFPSPPTTPSPAKTDQKKRSKFLPLFTNIEETLYKNVY 185
35      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
36 QY 190 HKIDLFVHLAKVLTNYEKSNNVYKIKELNLTITQDKLADFKNNNFVGIADLSTDY 249
37      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 Db 186 NKIDYILNLAKINDCNVEDEAHVKTTKLSDLAIDDKIDILFNTNDFAIKRLIND 245
39      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 QY 250 NNNNLLTFELSTGMV--FENLAKTVLSNLDGNLQGMNLISHOQCVKROCPONGCFRL 307
41      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
42 Db 246 TKRMLGKLLSTGLVQIFPN---TIIKLIBEKFQDMLNISHOQCVKROCPNSGCFRL 302
43      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 QY 308 DEREBCKCLLYKQEGDKCVENPPTCENNNGGCDADAKCTEEDSGSNGKRTTCCTKPD 367
45      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 Db 303 DEREBCKCLLYKQEGDKCEENPPTCENNNGGCDADATCTEEDSGSSRKRTTCCTKPD 362
47      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48 QY 368 STPLDGIIFS 378
49      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 Db 363 STPLDGIIFS 373
51      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
52
53 RESULT 5
54 US-08-230-919-3
55 Sequence 3, Application US/08290919
56 Patent No. 5720959
57
58 GENERAL INFORMATION:
59 APPLICANT: HOLDER, ANTHONY A.
60 APPLICANT: BLACKMAN, MICHAEL J.
61 APPLICANT: CHAPPEL, JONATHAN A.
62 TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
63 TITLE OF INVENTION: VACCINE
64 NUMBER OF SEQUENCES: 19
65 CORRESPONDENCE ADDRESSES:
66 ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
67 STREET: 1100 NEW YORK AVENUE, N.W.
68 CITY: WASHINGTON
69 STATE: D.C.
70 COUNTRY: U.S.A.
71 ZIP: 20005-3918
72
73 COMPUTER READABLE FORM:
74 MEDIUM TYPE: floppy disk
75 COMPUTER: IBM PC compatible

```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label=X
OTHER INFORMATION: /note="X" = M and N, or N"
US-08-290-919-3

Query Match      13.9%; Score 287; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      330 NNPCTENNNGCGDADAKTEEDSGSGNGKKTCECTKPDSPYLPFGIFCS 378
DB      2  NNPCTENNNGCGDADAKTEEDSGSGNGKKTCECTKPDSPYLPFGIFCS 50

RESULT 6
US-08-290-919-12
Sequence 12, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-290-919-12

Query Match      13.6%; Score 280; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      286 NISOHCVCVKKCPQNSGCFRHLDERECKCLINTKQSGDKCVENPNPT 333
DB      1  NISOHCVCVKKCPQNSGCFRHLDERECKCLINTKQSGDKCVENPNPT 48

RESULT 7
US-08-290-919-1
Sequence 1, Application US/08290919
Patent No. 5720959
GENERAL INFORMATION:
APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,919
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9203821.5
FILING DATE: 22-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00367
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
```

```

: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 1
: OTHER INFORMATION: /label= X
: OTHER INFORMATION: /note= "X" - M and N, or "N"
US-08-290-919-1

Query Match 13.3%; Score 274; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 3,2e-14;
Matches 47; Conservative 0; Mismatches 0; Indels

QY 287 ISQHCVKQKQPPNSGCFRHLDERECKCLLNTYKQEGDKCVENPNPT 333
|||||
Db 2 ISQHCVKQKQPPNSGCFRHLDERECKCLLNTYKQEGDKCVENPNPT 48

RESULT 8
US-08-290-919-2
: Sequence 2, Application US/08290919
: Patent No. 5720959
: GENERAL INFORMATION:
: APPLICANT: HOLDER, ANTHONY A.
: APPLICANT: BLACKMAN, MICHAEL J.
: APPLICANT: CHAPPEL, JONATHAN A.
: TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
: TITLE OF INVENTION: VACCINE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
: STREET: 1100 NEW YORK AVENUE, N.Y.
: CITY: WASHINGTON
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005-3918
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/290,919
: FILING DATE: 04-OCT-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9203821.5
: FILING DATE: 22-FEB-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB93/00367
: FILING DATE: 22-FEB-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: KOKULIS, PAUL N.
: REGISTRATION NUMBER: 16,773
: REFERENCE/DOCKET NUMBER: 212242/HCM/MLT/6BC8/
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 861-3000
: TELEFAX: (202) 822-0944
: TELEX: 6714627 CUSH
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 48 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 1
: OTHER INFORMATION: /label= X
: OTHER INFORMATION: /note= "X" - M and N, or "N"

```

Query Match	13.1%	Score 271;	DB 1;	Length 48;
Best Local Similarity	97.9%	Pred. No. 5,4e-14;		
Matches 46;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
<p>US-08-290-919-2</p> <p>Query Match 13.1%; Score 271; DB 1; Length 48;</p> <p>Best Local Similarity 97.9%; Pred. No. 5,4e-14;</p> <p>Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;</p>				
<p>US-08-290-919-4</p> <p>Sequence 4, Application US/08290919</p> <p>Patent No. 5720959</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: HOLDER, ANTHONY A.</p> <p>APPLICANT: BLACKMAN, MICHAEL J.</p> <p>TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA</p> <p>NUMBER OF INVENTION: VACCINE</p> <p>NUMBER OF SEQUENCES: 19</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.</p> <p>STREET: 1100 NEW YORK AVENUE, N.W.</p> <p>CITY: WASHINGTON</p> <p>STATE: D.C.</p> <p>COUNTRY: U.S.A.</p> <p>ZIP: 20005-3918</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent in Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/290,919</p> <p>FILING DATE: 04-OCT-1994</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: GB 9203821.5</p> <p>FILING DATE: 22-FEB-1992</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: PCT/GB93/00367</p> <p>FILING DATE: 22-FEB-1993</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: KOKULIS, PAUL N.</p> <p>REGISTRATION NUMBER: 16,773</p> <p>REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (202) 861-3000</p> <p>TELEFAX: (202) 822-0944</p> <p>TELEX: 671462Z CUSH</p> <p>INFORMATION FOR SEQ ID NO: 4:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 53 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>FEATURE:</p> <p>NAME/KEY: Modified-site</p> <p>LOCATION: 1</p> <p>OTHER INFORMATION: /label-X</p> <p>OTHER INFORMATION: /note="X" - M and N, or N"</p> <p>US-08-290-919-4</p>				
<p>Query Match 13.0%; Score 268; DB 1; Length 53;</p> <p>Best Local Similarity 93.9%; Pred. No. 1e-13;</p> <p>Matches 46; Conservative 1; Mismatches 2; Indels 0; Gaps 0;</p>				
<p>US-08-290-919-4</p> <p>Sequence 4, Application US/08290919</p> <p>Patent No. 5720959</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: HOLDER, ANTHONY A.</p> <p>APPLICANT: BLACKMAN, MICHAEL J.</p> <p>TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA</p> <p>NUMBER OF INVENTION: VACCINE</p> <p>NUMBER OF SEQUENCES: 19</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.</p> <p>STREET: 1100 NEW YORK AVENUE, N.W.</p> <p>CITY: WASHINGTON</p> <p>STATE: D.C.</p> <p>COUNTRY: U.S.A.</p> <p>ZIP: 20005-3918</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent in Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/290,919</p> <p>FILING DATE: 04-OCT-1994</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: GB 9203821.5</p> <p>FILING DATE: 22-FEB-1992</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: PCT/GB93/00367</p> <p>FILING DATE: 22-FEB-1993</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: KOKULIS, PAUL N.</p> <p>REGISTRATION NUMBER: 16,773</p> <p>REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (202) 861-3000</p> <p>TELEFAX: (202) 822-0944</p> <p>TELEX: 671462Z CUSH</p> <p>INFORMATION FOR SEQ ID NO: 4:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 53 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>FEATURE:</p> <p>NAME/KEY: Modified-site</p> <p>LOCATION: 1</p> <p>OTHER INFORMATION: /label-X</p> <p>OTHER INFORMATION: /note="X" - M and N, or N"</p> <p>US-08-290-919-4</p>				
<p>Query Match 13.0%; Score 268; DB 1; Length 53;</p> <p>Best Local Similarity 93.9%; Pred. No. 1e-13;</p> <p>Matches 46; Conservative 1; Mismatches 2; Indels 0; Gaps 0;</p>				
<p>US-08-290-919-4</p> <p>Sequence 4, Application US/08290919</p> <p>Patent No. 5720959</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: HOLDER, ANTHONY A.</p> <p>APPLICANT: BLACKMAN, MICHAEL J.</p> <p>TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA</p> <p>NUMBER OF INVENTION: VACCINE</p> <p>NUMBER OF SEQUENCES: 19</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.</p> <p>STREET: 1100 NEW YORK AVENUE, N.W.</p> <p>CITY: WASHINGTON</p> <p>STATE: D.C.</p> <p>COUNTRY: U.S.A.</p> <p>ZIP: 20005-3918</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent in Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/290,919</p> <p>FILING DATE: 04-OCT-1994</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: GB 9203821.5</p> <p>FILING DATE: 22-FEB-1992</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: PCT/GB93/00367</p> <p>FILING DATE: 22-FEB-1993</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: KOKULIS, PAUL N.</p> <p>REGISTRATION NUMBER: 16,773</p> <p>REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (202) 861-3000</p> <p>TELEFAX: (202) 822-0944</p> <p>TELEX: 671462Z CUSH</p> <p>INFORMATION FOR SEQ ID NO: 4:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 53 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>FEATURE:</p> <p>NAME/KEY: Modified-site</p> <p>LOCATION: 1</p> <p>OTHER INFORMATION: /label-X</p> <p>OTHER INFORMATION: /note="X" - M and N, or N"</p> <p>US-08-290-919-4</p>				
<p>Query Match 13.0%; Score 268; DB 1; Length 53;</p> <p>Best Local Similarity 93.9%; Pred. No. 1e-13;</p> <p>Matches 46; Conservative 1; Mismatches 2; Indels 0; Gaps 0;</p>				
<p>US-08-290-919-4</p> <p>Sequence 4, Application US/08290919</p> <p>Patent No. 5720959</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: HOLDER, ANTHONY A.</p> <p>APPLICANT: BLACKMAN, MICHAEL J.</p> <p>TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA</p> <p>NUMBER OF INVENTION: VACCINE</p> <p>NUMBER OF SEQUENCES: 19</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.</p> <p>STREET: 1100 NEW YORK AVENUE, N.W.</p> <p>CITY: WASHINGTON</p> <p>STATE: D.C.</p> <p>COUNTRY: U.S.A.</p> <p>ZIP: 20005-3918</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent in Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/290,919</p> <p>FILING DATE: 04-OCT-1994</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: GB 9203821.5</p> <p>FILING DATE: 22-FEB-1992</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: PCT/GB93/00367</p> <p>FILING DATE: 22-FEB-1993</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: KOKULIS, PAUL N.</p> <p>REGISTRATION NUMBER: 16,773</p> <p>REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (202) 861-3000</p> <p>TELEFAX: (202) 822-0944</p> <p>TELEX: 671462Z CUSH</p> <p>INFORMATION FOR SEQ ID NO: 4:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 53 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>FEATURE:</p> <p>NAME/KEY: Modified-site</p> <p>LOCATION: 1</p> <p>OTHER INFORMATION: /label-X</p> <p>OTHER INFORMATION: /note="X" - M and N, or N"</p> <p>US-08-290-919-4</p>				
<p>Query Match 13.0%; Score 268; DB 1; Length 53;</p> <p>Best Local Similarity 93.9%; Pred. No. 1e-13;</p> <p>Matches 46; Conservative 1; Mismatches 2; Indels 0; Gaps 0;</p>				
<p>US-08-290-919-4</p> <p>Sequence 4, Application US/08290919</p>				

RESULT 10

US-08-290-919-11

Sequence 11, Application US/08290919
Patent No. 5720959

GENERAL INFORMATION:

APPLICANT: HOLDER, ANTHONY A.
APPLICANT: BLACKMAN, MICHAEL J.
APPLICANT: CHAPPEL, JONATHAN A.
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO MALARIA
NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,919

FILING DATE: 04-OCT-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9203821.5

FILING DATE: 22-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB93/00367

FILING DATE: 22-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: KOKULIS, PAUL N.

REGISTRATION NUMBER: 16,773

REFERENCE/DOCKET NUMBER: 212242/HCM/MJL/6BC8/

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-290-919-11

Query Match

Best Local Similarity 12.3%; Score 253.5; DB 1; Length 106;

Matches 49; Conservative 17; Mismatches 34; Indels 5; Gaps 4;

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

TITLE OF INVENTION: Expression and Use of Human Fibroblast
Growth Factor Receptor

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/640,029

FILING DATE: 19910111

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: CH-165

TELEPHONE: 510-601-2708

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 816 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-640-029-1

Query Match

Best Local Similarity 7.08%; Score 144; DB 1; Length 816;

Matches 77; Conservative 40; Mismatches 94; Indels 100; Gaps 16;

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

TITLE OF INVENTION: Expression and Use of Human Fibroblast
Growth Factor Receptor

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/640,029

FILING DATE: 19910111

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: CH-165

TELEPHONE: 510-601-2708

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 816 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-640-029-1

Query Match

Best Local Similarity 24.88%; Score 144; DB 1; Length 816;

Matches 77; Conservative 40; Mismatches 94; Indels 100; Gaps 16;

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 4669
;; LENGTH: 384
;; TYPE: PRF
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4669

Query Match 6.9%; Score 141.5; DB 4; Length 384;
Best Local Similarity 27.8%; Pred. No. 0.0042;
Matches 69; Conservative 35; Mismatches 81; Indels 63; Gaps 13;

QY 43 VILKPLAGVRSIKKQLENNVMTFNVNKKDLNSFKNREKFNKLVESDLIPYDUTSS 102
DB 42 VVLYLSDVSKSL-----NKKTYAVALK-----NPFVYLESKKKCYKRVSPD 85
QY 103 NTV-----VKDPYKF---LNKEKRDKFLSS-----YNYIKSDITDINFAND 141
DB 86 NVDYDILRMKKTFEYENLVSYHREKKSISPKTINILMTVVSNFYDILYRSKLDVNF-YD 144
QY 142 VLGYYKILSEKYSKSDLSIKKIYND-----KQGENEKYLPPLNIET---LYTVMHKIDL 194
DB 145 FHMESKYSKKKYSFHMHNKNDYRFLKNILKYEPRKKIEVLNNAVKKILLEANIRDK 204
QY 195 FVHLEAKVLYNTEKSNVEKIKELNYKTIODKLADKKNNNPYGIADLSTDYHNNL 254
DB 205 FLIQ-----LYETG---LRIGEVLSLR-IDDIKFDPRKGHIV---LKNRFNDNG- 248
QY 255 LTKFLSTG 262
DB 249 --TYLKTG 254

RESULT 13

US-07-640-029-2
; Sequence 2, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chilton Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 817 amino acids
;; TYPE: AMINO ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-07-640-029-2

Query Match 6.9%; Score 141.5; DB 1; Length 817;
Best Local Similarity 24.4%; Pred. No. 0.011;
Matches 76; Conservative 41; Mismatches 93; Indels 101; Gaps 16;

QY 1 MMSWRCLEWAVLVATLTCTAGAAVT-----PSYIDNIL----- 34
DB 1 MMSWRCLEWAVLVATLTCTARSPFLPQAPMGAPVESEFLVHPGDLQLRCLRLD 60
QY 35 -----SKIE-NREYVLYLPL-AGVRSKQKQLENNVMT-FNVNVD 73
DB 61 VOSIMWLKRGVOLAASNRTIRIGEEYVODSVPAOSGLACVTSFSGSDTYTFSVNSD 120
QY 74 IINSFKNREKFNKLVESDLIPYKDLSSNYVVKDPYKFLNKEKRDKFL----- 122
DB 121 ALPSSDDDDDDSSSE---KETDNTKFNRPVAPY-WTSEKMEKILHAPAKTYKF 176
QY 123 -----SSYNYIKSDITDINFAND-VLGYYKILSEKYSKSDLSIKKIYINDROGENE 172
DB 177 KCPSSGTPMPTLRMLNKEK-----FKPDHRIIGYKRVATWGLINDSV---VPSDKG-- 227
QY 173 KTLPLNNTIELYKYNHKLIDLVH-----LEAKV-LNTYIE-KSNVEYKIR----- 218
DB 228 ----YTCIYENEGSINHTYQLDVVERSPHRLDAGLPANNTVALGVSVEEMCVYSDP 283
QY 219 --ELNYLTKTIQ 227
DB 284 OPHIQMLKHIE 294

RESULT 14

US-08-166-717D-6
; Sequence 6, Application US/08166717D
; Patent No. 5789182
; GENERAL INFORMATION:
; APPLICANT: Yayon, Avner
; APPLICANT: Ornitz, David M.
; APPLICANT: Klagsbrun, Michael
; APPLICANT: Leder, Philip
; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
; TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: Wordperfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,717D
; FILING DATE: 12/14/93
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 07/631,717
; FILING DATE: 12/20/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Kristina Bieker-Brady
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 00383/017002

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-4123
TELEFAX: (617) 723-8962
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 820
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-166-717D-6

Query Match 6.9%; Score 141.5; DB 1; Length 820;
Best Local Similarity 24.8%; Pred. No. 0.011;
Matches 77; Conservative 40; Mismatches 91; Indels 103; Gaps 17;

QY 1 MMSKCLLFMAVLTATICTAGAAVT-----PSVIDNLT----- 34
DB 1 MMSKCLLFMAVLTATICTARAPPTLPDAQPMGAVPEVESLLVHPGDLQLRCRLDD 60
QY 35 -----SKLEN-----EYEVLYLKPL-AGVYRSLKRLQLENNVMT-FNVNPKD 73
DB 61 VQSIWXXKDGQVLYESNRTRTGEVEVRDSIPADSGLYACVTSPPSGSDTYTFSVNVSD 120
QY 74 ILNSRFNKRNFKNVLESDDLIPYKDLTSSNTVVDPKYFLNKKERDKFL----- 122
DB 121 ALPSEDDDDHDSSESE---KETDNTKPNPVA--PY-WTSPKMEKKLHVPAPAKTVKF 174
QY 123 -----SSYNIKSIDIDTINFAND-VLGYYKILSEKYSKSDLSIKKYINDKOGENE 172
DB 175 KCPSSGPNPTLRKLKNGKE---FKPDHRIGGYKRVATWSIIMDSV---VPSDKGN-- 225
QY 173 KYLPFLNNIETLYKTVNKKIDLFYIH-----LEAVY-LNTYVE-KSNVEVKIK----- 218
DB 226 ---YTCIVENEYSINHNTYQLDVVERSPHRPILOAGLXANKYVALGSLNVEFMCKVYSDX 281
QY 219 --ELNVLKTIQ 227
DB 282 XPHIOMLKHIIE 292

RESULT 15
US-07-997-133-1
Sequence 1, Application US/07997133
Patent No. 5288855
GENERAL INFORMATION:
APPLICANT: Bergonzoni, Laura
APPLICANT: Mazue, Guy
APPLICANT: Isacchi, Antonella
APPLICANT: Roncucci, Romeo
APPLICANT: Sarmiento, Paolo
TITLE OF INVENTION: Extracellular Form of the Human
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 175 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/997,133
FILING DATE: 28-DEC-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/642,755
FILING DATE: 18-JAN-1991

ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5288855man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 769-226-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-997-133-1

Query Match 6.9%; Score 141.5; DB 1; Length 822;
Best Local Similarity 24.4%; Pred. No. 0.011;
Matches 76; Conservative 41; Mismatches 93; Indels 101; Gaps 16;

QY 1 MMSKCLLFMAVLTATICTAGAAVT-----PSVIDNLT----- 34
DB 1 MMSKCLLFMAVLTATICTARAPPTLPDAQPMGAVPEVESLLVHPGDLQLRCRLDD 60
QY 35 -----SKIE-NEYEVLYLKPL-AGVYRSLKRLQLENNVMT-FNVNPKD 73
DB 61 VQSIWMLRDGQVLAETNRTTRTGEVEVQDSVPADSGLYACVTSPPSGSDTYTFSVNVSD 120
QY 74 ILNSRFNKRNFKNVLESDDLIPYKDLTSSNTVVDPKYFLNKKERDKFL----- 122
DB 121 ALPSEDDDDHDSSESE---KETDNTKPNMPVAPY-WTSPKMEKKLHVPAPAKTVKF 176
QY 123 -----SSYNIKSIDIDTINFAND-VLGYYKILSEKYSKSDLSIKKYINDKOGENE 172
DB 177 KCPSSGPNPTLRKLKNGKE---FKPDHRIGGYKRVATWSIIMDSV---VPSDKGN-- 227
QY 173 KYLPFLNNIETLYKTVNKKIDLFYIH-----LEAVY-LNTYVE-KSNVEVKIK----- 218
DB 228 ---YTCIVENEYSINHNTYQLDVVERSPHRPILOAGLIPANKYVALGSLNVEFMCKVYSDP 283
QY 219 --ELNVLKTIQ 227
DB 284 QPHIOMLKHIIE 294

Search completed: June 11, 2003, 16:09:49
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:06:00 ; Search time 22 Seconds
(without alignments)
1802.012 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSMKCLFWAVLVATITCT.....KPDSTPLFDGIFCSHHHHH 384

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2063	100.0	384	9	US-10-062-809-8 Sequence 8, Appl1
2	1879	91.1	378	9	US-10-087-464-11 Sequence 11, Appl1
3	1879	91.1	1639	9	US-10-087-464-10 Sequence 10, Appl1
4	1874	90.8	376	9	US-09-978-756-2 Sequence 2, Appl1
5	1874	90.8	376	9	US-10-087-464-34 Sequence 34, Appl1
6	1719.5	83.3	377	9	US-10-062-809-5 Sequence 5, Appl1
7	1658.5	80.4	375	9	US-10-062-809-4 Sequence 4, Appl1
8	1347	65.3	402	9	US-10-098-514-14 Sequence 14, Appl1
9	1347	65.3	402	9	US-10-062-809-16 Sequence 16, Appl1
10	11310	63.5	396	9	US-10-062-809-18 Sequence 18, Appl1
11	1198.5	58.1	383	9	US-10-098-514-4 Sequence 4, Appl1
12	1197.5	57.8	383	9	US-10-098-514-2 Sequence 2, Appl1
13	1191.5	57.8	394	9	US-10-062-809-2 Sequence 2, Appl1
14	1161.5	56.3	394	9	US-09-978-756-3 Sequence 3, Appl1
15	1156.5	56.1	394	9	US-10-062-809-3 Sequence 3, Appl1
16	742.5	36.0	380	10	US-09-134-333-13 Sequence 13, Appl1
17	738.5	35.8	380	10	US-09-134-333-12 Sequence 12, Appl1
18	676	32.8	379	10	US-09-134-333-11 Sequence 11, Appl1
19	542	26.3	96	9	US-09-978-736-1 Sequence 1, Appl1

20	542	26.3	114	9	US-10-087-464-35 Sequence 35, Appl1
21	539	26.1	95	10	US-09-134-333-2 Sequence 2, Appl1
22	539	26.1	108	10	US-09-134-333-10 Sequence 10, Appl1
23	539	26.1	116	10	US-09-134-333-5 Sequence 5, Appl1
24	539	26.1	127	10	US-09-134-333-8 Sequence 8, Appl1
25	426.5	20.7	281	10	US-09-134-333-14 Sequence 14, Appl1
26	170	8.2	1394	9	US-09-842-930A-25 Sequence 25, Appl1
27	139.5	6.8	822	9	US-09-757-415A-2 Sequence 2, Appl1
28	136.5	6.6	493	9	US-10-234-432-24 Sequence 24, Appl1
29	136.5	6.6	893	9	US-10-234-432-86 Sequence 86, Appl1
30	128	6.2	497	9	US-09-820-843A-32 Sequence 32, Appl1
31	128	6.2	861	9	US-09-820-843A-109 Sequence 109, Appl1
32	124	6.0	358	10	US-09-861-451A-50 Sequence 50, Appl1
33	123.5	6.0	1143	10	US-09-924-154-14 Sequence 14, Appl1
34	123	6.0	471	10	US-09-925-302-788 Sequence 788, Appl1
35	122.5	5.9	996	10	US-09-815-242-5251 Sequence 5251, Appl1
36	122.5	5.9	1009	10	US-09-815-242-12141 Sequence 12141, Appl1
37	121	5.9	334	9	US-10-087-464-33 Sequence 33, Appl1
38	121	5.9	360	9	US-10-087-464-12 Sequence 12, Appl1
39	121	5.9	1431	9	US-09-842-930A-2 Sequence 2, Appl1
40	120.5	5.8	778	9	US-09-908-193-8 Sequence 8, Appl1
41	120.5	5.8	779	9	US-09-908-193-10 Sequence 10, Appl1
42	120.5	5.8	841	10	US-09-861-451A-30 Sequence 30, Appl1
43	120.5	5.8	1086	10	US-09-924-154-15 Sequence 15, Appl1
44	120	5.8	621	10	US-09-996-620-6 Sequence 6, Appl1
45	120	5.8	754	9	US-09-908-193-32 Sequence 32, Appl1

ALIGNMENTS

RESULT 1
US-10-062-809-8
; Sequence 8, Application US/10062809
; Publication No. US20030100106A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra P.
; APPLICANT: Hashimoto, Ann
; APPLICANT: Hashimoto, Ann
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM FALCIPARUM VACCINE
; FILE REFERENCE: A-67984-1/REF/TAI/NBC
; CURRENT APPLICATION NUMBER: US/10/062, 809
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/500, 376
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 60/266, 281
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-062-809-8

Query Match 100.0%; Score 2063; DB 9; Length 384;
Best Local Similarity 100.0%; Pred. No. 9, 6e-125;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSMKCLFWAVLVATITCTAGAAVPSYIDNLSIKENYEYLYTKPLAGYRSIKKOL 60
DB 1 MMSMKCLFWAVLVATITCTAGAAVPSYIDNLSIKENYEYLYTKPLAGYRSIKKOL 60
QY 61 ENNVMTFNNVNDKILSRFNKRNPNFNVESDILPKDITSSNYVKKDKYKFLNKKRKK 120
DB 61 ENNVMTFNNVNDKILSRFNKRNPNFNVESDILPKDITSSNYVKKDKYKFLNKKRKK 120
QY 121 FLSSTNYIKDSIDTIDNFANDVLYGYKKIISSEKYSKSLDSIKKYINDKQENKYLPLFNN 180
DB 121 FLSSTNYIKDSIDTIDNFANDVLYGYKKIISSEKYSKSLDSIKKYINDKQENKYLPLFNN 180
QY 181 IETLYTVNKKIDLEFYIHLEAKVLYNTTYSKSNVEVKKIKELNLYTKITQDLADPKNNNFV 240
DB 181 IETLYTVNKKIDLEFYIHLEAKVLYNTTYSKSNVEVKKIKELNLYTKITQDLADPKNNNFV 240

Db 181 IETKYVNHKIDLFVHLEAKVINYTEKSNVEVKIKELNYLKTIOCKLADFKKNNNFV 240
Qy 241 GIADLSTDYNNHNLITKELSTGMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPON 300
Db 241 GIADLSTDYNNHNLITKELSTGMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPON 300
Qy 301 SGCFRHLDERECKCLNYKOEKDCVBNPPTCNENNGGCDADAKTEEDSGSNKKKIT 360
Db 301 SGCFRHLDERECKCLNYKOEKDCVBNPPTCNENNGGCDADAKTEEDSGSNKKKIT 360
Qy 361 CECTKPDSPLEDFGIFCSHHHHH 384
Db 361 CECTKPDSPLEDFGIFCSHHHHH 384

RESULT 2

US-10-087-464-11
Sequence 11, Application US/10087464
Publication No. US20030059436A1

GENERAL INFORMATION:

APPLICANT: Chishtli, Alhar
APPLICANT: Oh, Steven
APPLICANT: Liu, David
APPLICANT: Goel, Vikas
APPLICANT: Li, Xuerong
TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
FILE REFERENCE: S1237/7019
CURRENT FILING DATE: 2002-03-01
CURRENT APPLICATION NUMBER: US/10/087,464
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 06/272,930
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 378
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-087-464-11

Query Match 91.1%; Score 1879; DB 9; Length 378;
Best Local Similarity 99.4%; Pred. No. 5,7e-113;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 22 GAATPSYIDNIIKIENEYEVLYLKLPLAGYRSLSKQLENNVTFVNNVADIIINSRPNK 81
Db 1 GEATPSYIDNIIKIENEYEVLYLKLPLAGYRSLSKQLENNVTFVNNVADIIINSRPNK 60
Qy 82 RENFNKLVESDLIPYKDLTSSNYVVKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 141
Db 61 RENFNKLVESDLIPYKDLTSSNYVVKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 120
Qy 142 VLGYYKILSEKYSKDDLSIKKYINDKOGENEKYLPLNNIETLYKTVNHKIDLFVHLEA 201
Db 121 VLGYYKILSEKYSKDDLSIKKYINDKOGENEKYLPLNNIETLYKTVNHKIDLFVHLEA 180
Qy 202 KVLNYYTEKSNVEVKIKELNLTOKLADFKKNNNFVGIADISTYNNHNLITKFLST 261
Db 181 KVLNYYTEKSNVEVKIKELNLTOKLADFKKNNNFVGIADISTYNNHNLITKFLST 240
Qy 262 GMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPONSGCFRHLDERECKCLNYKQ 321
Db 241 GMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPONSGCFRHLDERECKCLNYKQ 300
Qy 322 EGDRCVENPPTCNENNGGCDADAKTEEDSGSNKKKITCECTKPDSPLEDFGIFCS 378
Db 301 EGDRCVENPPTCNENNGGCDADAKTEEDSGSNKKKITCECTKPDSPLEDFGIFCS 357

RESULT 3
US-10-087-464-10
Sequence 10, Application US/10087464
Publication No. US20030059436A1
GENERAL INFORMATION:

APPLICANT: Chishtli, Alhar
APPLICANT: Oh, Steven
APPLICANT: Liu, David
APPLICANT: Goel, Vikas
APPLICANT: Li, Xuerong
TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses There
FILE REFERENCE: S1237/7019
CURRENT FILING DATE: 2002-03-01
CURRENT APPLICATION NUMBER: US/10/087,464
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 06/272,930
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 1639
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-087-464-10

Query Match 91.1%; Score 1879; DB 9; Length 1639;
Best Local Similarity 99.4%; Pred. No. 3e-112;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 22 GAATPSYIDNIIKIENEYEVLYLKLPLAGYRSLSKQLENNVTFVNNVADIIINSRPNK 81
Db 1262 GEATPSYIDNIIKIENEYEVLYLKLPLAGYRSLSKQLENNVTFVNNVADIIINSRPNK 1321
Qy 82 RENFNKLVESDLIPYKDLTSSNYVVKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 141
Db 1222 RENFNKLVESDLIPYKDLTSSNYVVKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 1381
Qy 142 VLGYYKILSEKYSKDDLSIKKYINDKOGENEKYLPLNNIETLYKTVNHKIDLFVHLEA 201
Db 1382 VLGYYKILSEKYSKDDLSIKKYINDKOGENEKYLPLNNIETLYKTVNHKIDLFVHLEA 1441
Qy 202 KVLNYYTEKSNVEVKIKELNLTOKLADFKKNNNFVGIADISTYNNHNLITKFLST 261
Db 1442 KVLNYYTEKSNVEVKIKELNLTOKLADFKKNNNFVGIADISTYNNHNLITKFLST 1501
Qy 262 GMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPONSGCFRHLDERECKCLNYKQ 321
Db 1502 GMYFENLAKTVLSNLDGNLQGMNLISOHOCVKKQCPONSGCFRHLDERECKCLNYKQ 1561
Qy 322 EGDRCVENPPTCNENNGGCDADAKTEEDSGSNKKKITCECTKPDSPLEDFGIFCS 378
Db 1562 EGDRCVENPPTCNENNGGCDADAKTEEDSGSNKKKITCECTKPDSPLEDFGIFCS 1618

RESULT 4

US-09-978-756-2
Sequence 2, Application US/09978756
Patent No. US20020160017A1

GENERAL INFORMATION:

APPLICANT: Holder, Anthony
APPLICANT: Birdsell, Betty
APPLICANT: Feeney, James
APPLICANT: Morgan, William
APPLICANT: Syed, Shabih
TITLE OF INVENTION: Malaria Vaccine
FILE REFERENCE: 18396/1005
CURRENT FILING DATE: 2001-10-16
CURRENT APPLICATION NUMBER: US/09/978,756
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: 09/311,817
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 2,271,451
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 9909072.2
PRIOR FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2

```

; LENGTH: 376
; TYPE: PRF
; ORGANISM: Plasmodium falciparum
US-09-978-756-2

```

Query Match	90.8%	Score 1874;	DB 9;	Length 376;
Best Local Similarity	99.7%	Pred. NO. 1.2e-112;		
Matches 354; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	24	AVPVSVIDNLSIKIENEVEVYLKPLAGVYSLSKLQKLNWTFVWVVKDILNSPKNRE	83
Db	1	AVPVSVIDNLSIKIENEVEVYLKPLAGVYSLSKLQKLNWTFVWVVKDILNSPKNRE	60
QY	84	NEKRVLESDDLPRYDITSSNVVDPRYELKKEKRDFLSSVNTKQSDIDIDINRANVYL	143
Db	61	NEKRVLESDDLPRYDITSSNVVDPRYELKKEKRDFLSSVNTKQSDIDIDINRANVYL	120
QY	144	GYKRLSEKYSKSDLSIRKYINDKQGENEKLPLPLNINTEPLYKYVNHKIDLEFVHLEAKV	203
Db	121	GYKRLSEKYSKSDLSIRKYINDKQGENEKLPLPLNINTEPLYKYVNHKIDLEFVHLEAKV	180
QY	204	LNRYEESNVSVKIKELNYLKTIDOKLADFKNNNFVGIADLSTYNNNNLLTKRLSLGGM	263
Db	181	LNRYEESNVSVKIKELNYLKTIDOKLADFKNNNFVGIADLSTYNNNNLLTKRLSLGGM	240
QY	264	VFENILAKTVLSNLDGNGIQMLNITSQHCVKQCPQNSGCFRHLDEREBCKILMYKQEG	323
Db	241	VFENILAKTVLSNLDGNGIQMLNITSQHCVKQCPQNSGCFRHLDEREBCKILMYKQEG	300
QY	324	DKCYENPRPTCENNNNGGCDAAKCTEBSGSGSNKKITTECBKTPDSYPLPFDGIFCS	378
Db	301	DKCYENPRPTCENNNNGGCDAAKCTEBSGSGSNKKITTECBKTPDSYPLPFDGIFCS	355

RESULT 5
US-10-087-464-34

```

? Sequence 34 Application US/10087464
? Publication No. US20030059436A1
? GENERAL INFORMATION:
? APPLICANT: Chishti, Alhar
? APPLICANT: Oh, Steven
? APPLICANT: Liu, David
? APPLICANT: Goel, Vikas
? APPLICANT: Li, Xuerong
? TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
? FILE REFERENCE: S1237/77019
? CURRENT APPLICATION NUMBER: US/10/087,464
? CURRENT FILING DATE: 2002-03-01
? PRIOR APPLICATION NUMBER: US 06/272,930
? PRIOR FILING DATE: 2001-03-02
? NUMBER OF SEQ ID NOS: 59
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 34
? LENGTH: 376
? TYPE: PRT
? ORGANISM: Plasmodium falciparum
? US-10-087-464-34

```

Query Match	90.8%	Score 1874	DB 9	Length 376
Best Local Similarity	99.7%	Pred. No. 1.2e-112		
Matches 354	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	24	AVTPEVIDNIIISKIENEYEVYLKPLACVYRSISLKOLENNVATFVNVVADIDINSRPNKRE	83	
Db	1	AVTPEVIDNIIISKIENEYEVYLKPLACVYRSISLKOLENNVATFVNVVADIDINSRPNKRE	60	
QY	84	NFKNVLESDDLIPYKDLTSSNYYVADPYFKLNKKRDKFLSSYNYTKDSITDIDINPANDVL	143	
Db	61	NFKNVLESDDLIPYKDLTSSNYYVADPYFKLNKKRDKFLSSYNYTKDSITDIDINPANDVL	120	
QY	144	GYVYIILSKRYSDSDSIKKYINDKQGENEKKIPLFANNIETLYKYTNHKTIDLFVHILKARY	203	
Db	121	GYVYIILSKRYSDSDSIKKYINDKQGENEKKIPLFANNIETLYKYTNVDKIDLFVHILKARY	180	

QY	204	LNTYKSNVEYIKELANTLKTLOKLDKFKNNHFGIADLSDYNNNLTFEISTGM	263
Db	181	LNTYKSNVEYIKELANTLKTLOKLDKFKNNHFGIADLSDYNNNLTFEISTGM	240
QY	264	VFENLAKTYVLSNLDJENLOGMNISOHOCVKQCPONSGCFPHLDRECKCLNTRYOEG	322
Db	241	VFENLAKTYVLSNLDJENLOGMNISOHOCVKQCPONSGCFPHLDRECKCLNTRYOEG	300
QY	324	DKCEVNPPTCCNNNGCCDAKACTEEDBSGSGKRTCECTRPDSYPLFDGIFCS	378
Db	301	DKCEVNPPTCCNNNGCCDAKACTEEDBSGSGKRTCECTRPDSYPLFDGIFCS	355

RESULT 6
US-10-062-809-5

```

: Sequence 5, Application US/1/0062809
: Publication No. US20030100106A1
: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra P.
: APPLICANT: Hashimoto, Ann
: APPLICANT: Hashimoto, Tami
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM FALCIPARUM VACCINANT
: FILE REFERENCE: A-67984-1/RFT/TAL/NBC
: CURRENT APPLICATION NUMBER: US/10/062, 809
: CURRENT FILING DATE: 2002-06-28
: PRIOR APPLICATION NUMBER: US 09/500, 376
: PRIOR FILING DATE: 2000-02-08
: PRIOR APPLICATION NUMBER: US 60/266, 281
: PRIOR FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 79
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 377
: TYPE: PRN
: ORGANISM: plasmodium falciparum
: US-10-062-809-5

```

Query Match	83.38;	Score 1719.5;	DB 9;	Length 377;
Best Local Similarity	93.08;	Pred. No. 9.1e-103;		
Matches 331;	Conservative 6;	Mismatches 18;	Indels 1;	Gaps 1

QY	2	AAMPVADNLSTKINENEYVLKPLAGYRSRLKQLENNWTFEYVVKDILNSRPKRE	83
Db	1	AATSVADNLSTKINENEYVLKPLAGYRSRLKQLENNWTFEYVVKDILNSRPKRE	60
QY	84	NFKNVLSDLLPYKDLTSSNYVKDPYKFLNKKDFLSSYNTIKDSIPTDINFADV	143
Db	61	NFKNVLSDLLPYKDLTSSNYVKDPYKFLNKKDFLSSYNTIKDSIPTDINFADV	120
QY	144	GYKFLSEKKSDLSIRKYINDKQGENKYLPFLNNIETLYKYVNHIDTFLYHLKAV	203
Db	121	GYKFLSEKKSDLSIRKYINDKQGENKYLPFLNNIETLYKYVNHIDTFLYHLKAV	180
QY	204	LNTTYEKSNEVAKIKELANTLKITIIOKLDAPFKNNNFVGIADLSTDYNNNLITFLSTGM	265
Db	181	LNTTYEKSNEVAKIKELANTLKITIIOKLDAPFKNNNFVGIADLSTDYNNNLITFLSTGM	240
QY	264	VFENILAKTVLSNLDENL - OGMNLSIOHOCVKKCPONSQSGRIHDEBECKCLANIKOE	322
Db	241	VFENILAKTVLSNLDENL - OGMNLSIOHOCVKKCPONSQSGRIHDEBECKCLANIKOE	300
QY	323	GDKCVENPNTCENNNGGCDADAKCTEEDSGSGNGKKITCECTKPDSPYLPFGITCS	378
Db	301	GDKCVENPNTCENNNGGCDADAKCTEEDSGSGNGKKITCECTKPDSPYLPFGITCS	356

RESULT 7
US-10-062-809-4

```

; Sequence 4, Application US/10062809
; Publication No. US20030100106A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra P.

```

```

: APPLICANT: Hashimoto, Ann
: APPLICANT: Nishimura, Tani
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM FALCIPARUM VACCINE
: FILE REFERENCE: A-67984-1/RET/TAL/NBC
: CURRENT APPLICATION NUMBER: US/10/062, 809
: PRIOR FILING DATE: 2002-06-28
: PRIOR APPLICATION NUMBER: US 09/500, 376
: PRIOR FILING DATE: 2000-02-08
: PRIOR APPLICATION NUMBER: US 60/266, 281
: PRIOR FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 79
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 375
: TYPE: PRT
: ORGANISM: Plasmodium falciparum
: US-10-062-809-4

Query Match      80.4% Score 1658.5; DB 9; Length 375;
Best Local Similarity 90.7%; Pred. No. 7.3e-99;
Matches 323; Conservative 7; Mismatches 23; Indels 3; Gaps 2;

QY 24 AVTPSVIDNITLSKTENEYEVLYLKLPLAGVYSLSKKOLENNVMTFNVVVKDILNSRFNKR 83
DB 1 AVTPSVIHNILSKTENEYEVLYLKLPLAGVYSLSKKOLENNVMTFNVVVKDILNSRFNKR 60
QY 84 NFKVAVESDLIPYDILSSNVVVKPYKFLNKKERKDFLSYNTIKOSIDTIDINPANDVL 143
DB 61 NFKVAVESDLIPYDILSSNVVVKPYKFLNKKERKDFLSYNTIKOSIDTIDINPANDVL 120
QY 144 GYVYILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKYVNHKIDLFVHLEAKV 203
DB 121 GYVYILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETLYKYVNHKIDLFVHLEAKV 180
QY 204 LNTYERSNVEVKIKELNLYKTIDKLADFERNKNNFVGADLSTDYNNHNNLTFFELSTGM 263
DB 181 LNTYERSNVE--IKELIYLTKTIDKLADFERNKNNFVGADLSTDYNNHNNLTFFELSTGM 238
QY 264 VFEMLATVLSNLDLGNL--QGMNLISQHCYKCKCPNSGCFRHLDERECKCLANTKOE 322
DB 239 VFEMLATVLSNLDLGNL--QGMNLISQHCYKCKCPNSGCFRHLDERECKCLANTKOE 298
QY 323 GDCVENVNPTCENNGCGDADACTEEDSGSGNKRTICECTKPSYPLFDGIFCS 378
DB 299 GSKCVENVNPTCENNGCGDADACTEEDSGSGNKRTICECTKPSYPLFDGIFCS 354

RESULT 8
US-10-098-514-14
: Sequence 14, Application US/10098514
: Publication No. US20020194648A1
: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra P
: APPLICANT: Christopher, David A
: APPLICANT: Vine, Benjamin
: APPLICANT: Su, Wei-Wen
: APPLICANT: Bugos, Robert
: TITLE OF INVENTION: PLASMIDIUM FALCIPARUM MEROZOITE SURFACE PROTEIN-1 MALARIA PRODUCT
: FILE REFERENCE: A-71339/RET/TAL/NBC
: CURRENT APPLICATION NUMBER: US/10/098, 514
: PRIOR FILING DATE: 2002-08-06
: PRIOR APPLICATION NUMBER: US 09/500, 376
: PRIOR FILING DATE: 2000-02-08
: PRIOR APPLICATION NUMBER: US 60/274, 599
: PRIOR FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 14
: LENGTH: 402
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:

```

```

: OTHER INFORMATION: synthetic
: US-10-098-514-14

Query Match      65.3% Score 1347; DB 9; Length 402;
Best Local Similarity 61.8%; Pred. No. 6.8e-79;
Matches 251; Conservative 55; Mismatches 74; Indels 26; Gaps 4;

QY 1 MMSWCKLLFMAVAVLTATCTAGAAVTPSVIDNITLSKTENEYEVLYLKLPLAGVYSLSKKOL 60
DB 1 MMSWCKLLFMAVAVLTATCTAGAAVTPSVIDNITLSKTENEYEVLYLKLPLAGVYSLSKKOL 57
QY 61 ENNVMTFNVVVKDILNSRFNKRNFKNVLESDDLIPYDILSSNVVVKPYKFLNKKERKDF 120
DB 58 ENNVMTFNVVVKDILNSRFNKRNFKNVLESDDLIPYDILSSNVVVKPYKFLNKKERKDF 117
QY 121 FLSSYNTIKOSIDTIDINPANDVLGYKILSEKYSKSDLSIKKYI----- 164
DB 118 LKSKYKIKESVENDINFAQGISYIEKVALAKYKODLESIKKYIKEKEKFPSSPTTPP 177
QY 165 ----NDKOGENEKYLPLNNIETLYKYVNHKIDLFVHLEAKVNLTYERSNVEVKIKEL 220
DB 178 SPATDEQKSKSKFLPELTNIETLYNNLVNKRIDYLLNLKAKINDCVNKEDEAHVKTITKL 237
QY 221 NYLKTIDKLADFERNKNNFVGADLSTDYNNHNNLTFFELSGVYFEMLATVLSNLDGN 280
DB 238 SOLAIDDKIDLFERNHDFDAIKRLINDDTKDMGLSTGLV--QNPPTIISKLEIGK 296
QY 281 LOGMLNISOHCYKCKCPNSGCFRHLDERECKCLANTYKOGDGCVENNPTCENNGG 340
DB 297 FQDMLNISOHCYKCKCPNSGCFRHLDERECKCLANTYKOGDGCVENNPTCENNGG 356
QY 341 CDADACTEEDSGSGNKRTICECTKPSYPLFDGIFCS--HHNNH 384
DB 357 CDADACTEEDSGSGNKRTICECTKPSYPLFDGIFCSPPHHNNH 402

RESULT 9
US-10-062-809-16
: Sequence 16, Application US/10062809
: Publication No. US20030100106A1
: GENERAL INFORMATION:
: APPLICANT: Chang, Sandra P
: APPLICANT: Hashimoto, Ann
: APPLICANT: Nishimura, Tani
: APPLICANT: Bugos, Robert
: TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMIDIUM FALCIPARUM VACCINE
: FILE REFERENCE: A-67984-1/RET/TAL/NBC
: CURRENT APPLICATION NUMBER: US/10/062, 809
: PRIOR FILING DATE: 2002-06-28
: PRIOR APPLICATION NUMBER: US 09/500, 376
: PRIOR FILING DATE: 2000-02-08
: PRIOR APPLICATION NUMBER: US 60/266, 281
: PRIOR FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 79
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 16
: LENGTH: 402
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: synthetic
: US-10-062-809-16

Query Match      65.3% Score 1347; DB 9; Length 402;
Best Local Similarity 61.8%; Pred. No. 6.8e-79;
Matches 251; Conservative 55; Mismatches 74; Indels 26; Gaps 4;

QY 1 MMSWCKLLFMAVAVLTATCTAGAAVTPSVIDNITLSKTENEYEVLYLKLPLAGVYSLSKKOL 60
DB 1 MMSWCKLLFMAVAVLTATCTAGAAVTPSVIDNITLSKTENEYEVLYLKLPLAGVYSLSKKOL 57
QY 61 ENNVMTFNVVVKDILNSRFNKRNFKNVLESDDLIPYDILSSNVVVKPYKFLNKKERKDF 120
DB 58 ENNVMTFNVVVKDILNSRFNKRNFKNVLESDDLIPYDILSSNVVVKPYKFLNKKERKDF 117

```



```

RESULT 12
US-10-098-514-2
; Sequence 2, Application US/10098514
; Publication No. US20020194648A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra P
; APPLICANT: Christopher, David A
; APPLICANT: Vine, Benjamin
; APPLICANT: Su, Wei-Wen
; APPLICANT: Bugos, Robert
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MEROZOITE SURFACE PROTEIN-1 MALARIA PRODUCE
; TITLE OF INVENTION: TRANSGENIC PLANTS
; FILE REFERENCE: A-71339/REF/TAL/NBC
; CURRENT APPLICATION NUMBER: US/10/098,514
; CURRENT FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: US 09/500,376
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 60/274,599
; PRIOR FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 383
; TYPE: prt
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: syathetic
; NAME/KEY: MISC.FEATURE
; LOCATION: (380)..(380)
; OTHER INFORMATION: "Xaa" at position 380 represents a stop codon
US-10-098-514-2

Query Match      58.0%; Score 1197.5; DB 9; Length 383;
Best Local Similarity 60.4%; Pred. No. 2.4e-69;
Matches 223; Conservative 53; Mismatches 72; Indels 21; Gaps 2;

QY 31 DNILSKTENEEVYLKPLAGYRSLSKQLENNVMTFNVNVKDIINSRKNRKNVLE 90
DB 9 DNILSGFENEVDVILKPLAGYRSLSKQLENNVMTFNVNVKDIINSRKNRKNVLE 68
QY 91 SOLIYKDLTSSNVVADPYKFLNKKRDKFLSSYNTIKSDIPIIDINPANDVLYKYLIS 150
DB 69 SDLMQFKHISNNEYIIEDSFRLNSEOKNTLTKSTKYIKESYENDIKFAQEGISYIEVYL 128
QY 151 EKYKSDLSIKKYI-----NDKOGENEKYLPLANNIETLYKTVNH 190
DB 129 AKYKDLDESIRKVIKEEKEKPPSPPTPPSPAKTDEQKESKFLPLTNIETLYNNLVN 188
QY 191 KIDLEVIHLEAVLNTYTEKSNEVEYKIKELNTLTIODKLADFKNNNFVGIAIDSTYN 250
DB 189 KIDLDYLNKAKINCINVEKDEAHVYIKTSLDKAIDKIDLFKNHNFDAIKKLINDT 248
QY 251 HNNLTFKSTGMEVFNLAITYLSMLDGNLQGMNLSOHQCVKQKQPCNSGCFRHIDE 310
DB 249 KDMGLKLSLSTGLV-QNPPNIIISKIISKGRQDMNINISOHQCVKQKQPCNSGCFRHIDE 307
QY 311 EECCKLNYKQEGDKCVENPPTCENNNGGCDADAKTEEDSGSGNGKKTCECTKPDSP 370
DB 308 EECCKLNYKQEGDKCVENPPTCENNNGGCDADAKTEEDSGSGNGKKTCECTKPDSP 367
QY 371 LFDGIFCSH 379
DB 368 LFDGIFCSH 376

```

```

; APPLICANT: Hashimoto, Ann
; APPLICANT: Nishimura, Tani
; TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM FALCIPARUM VACCINE
; FILE REFERENCE: A-67984-1/REF/TAL/NBC
; CURRENT APPLICATION NUMBER: US/10/062,809
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 09/500,376
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 60/266,281
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 394
; TYPE: prt
; ORGANISM: plasmodium falciparum
US-10-062-809-2

Query Match      57.8%; Score 1191.5; DB 9; Length 394;
Best Local Similarity 60.2%; Pred. No. 6e-69;
Matches 222; Conservative 54; Mismatches 72; Indels 21; Gaps 2;

QY 30 IDNLSKTEENEEVYLKPLAGYRSLSKQLENNVMTFNVNVKDIINSRKNRKNVLE 89
DB 6 MDNLSGFENEVDVILKPLAGYRSLSKQLENNVMTFNVNVKDIINSRKNRKNVLE 65
QY 90 ESDLIYKDLTSSNVVADPYKFLNKKRDKFLSSYNTIKSDIPIIDINPANDVLYKYL 149
DB 66 ESDLMQFKHISNNEYIIEDSFRLNSEOKNTLTKSTKYIKESYENDIKFAQEGISYIEVYL 125
QY 150 SEKYKSDLSIKKYI-----NDKOGENEKYLPLANNIETLYKTVN 189
DB 126 LAKYKDLDESIRKVIKEEKEKPPSPPTPPSPAKTDEQKESKFLPLTNIETLYNNLVN 185
QY 190 KIDLEVIHLEAVLNTYTEKSNEVEYKIKELNTLTIODKLADFKNNNFVGIAIDSTYN 249
DB 186 KIDLDYLNKAKINCINVEKDEAHVYIKTSLDKAIDKIDLFKNHNFDAIKKLINDT 245
QY 250 HNNLTFKSTGMEVFNLAITYLSMLDGNLQGMNLSOHQCVKQKQPCNSGCFRHIDE 309
DB 246 TKDMLGKLSLSTGLV-QNPPNIIISKIISKGRQDMNINISOHQCVKQKQPCNSGCFRHIDE 304
QY 310 REECCKLNYKQEGDKCVENPPTCENNNGGCDADAKTEEDSGSGNGKKTCECTKPDSP 369
DB 305 REECCKLNYKQEGDKCVENPPTCENNNGGCDADAKTEEDSGSGNGKKTCECTKPDSP 364
QY 370 PLFDGIFCS 378
DB 365 PLFDGIFCS 373

```

```

RESULT 13
US-10-062-809-2
; Sequence 2, Application US/10062809
; Publication No. US20030100106A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Sandra P.

```

```

RESULT 14
US-09-978-756-3
; Sequence 3, Application US/09978756
; Patent No. US20020160017A1
; GENERAL INFORMATION:
; APPLICANT: Holder, Anthony
; APPLICANT: Birdsall, Betty
; APPLICANT: Feeney, James
; APPLICANT: Morgan, William
; APPLICANT: Syed, Shabih
; TITLE OF INVENTION: Malaria Vaccine
; FILE REFERENCE: 18396/1005
; CURRENT APPLICATION NUMBER: US/09/978,756
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: PCF/GB00/01558
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 09/311,817
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 2,271,451
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 9909072.2
; PRIOR FILING DATE: 1999-04-20

```

NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 394
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-09-978-756-3

Query Match 56.3%; Score 1161.5; DB 9; Length 394;
Best Local Similarity 59.1%; Pred. No. 56-67;
Matches 218; Conservative 54; Mismatches 76; Indels 21; Gaps 2;

QY 30 IDNLSKIEEYEVLYLKLACGYRSLLKQLENNVTFENVVYKDIINSRPNREKFNVL 89
DB 6 MNIIISGFENEYDVLYIKLACGYRSLLKQLENNVTFENVVYKDIINSRPNREKFNVL 65
QY 90 ESDLIPIKDLTSSNVVVDPIYKFLNKKRDKFLSSYNTIKDSIDPIDINPANDVLYGYKTL 149
DB 66 ESDLIQFKHISSENEYIIEDSFKLNSSEQKNTLKSSTYKIKESVENDIKFAQGISYERY 125
QY 150 SEKYKSDLSIKRYI-----NDKQGENEKYLPFLNIIETLYKTVN 189
DB 126 LAKYKDDLESIKKVIKEKEKFPSSPPTPPSPAKTDEQKKESKFLPFLNIIETLYKTVN 185
QY 190 HKIDLFVHLKAKVLYTEKSNVEYKIKELNYLKTIDKLADFKNKNNFVGIADLSTDY 249
DB 186 NKIDYLLIMKAKINDCNVEKDEAHYKTKLSDIAIDKIDLFKNTNDFEAIKRLIND 245
QY 250 NNNMLTKFLSTGMV--FENLAKTVLSNLLDGNLQGMNISOHCYKQCPONGSGCFRHL 309
DB 246 TKKDMGLKSLSTGLV--QNPNTIISKLFQDMNISOHCYKQCPONGSGCFRHLDE 304
QY 310 REECKCLLNTKQEGDKCVENPPTCNENNGGCDADAKCTEEDSGSNKKTTCCTKPDY 369
DB 305 REECKCLLNTKQEGDKCVENPPTCNENNGGCDADATCTEEDSGSNKKTTCCTKPDY 364
QY 370 PLFDGIFCS 378
DB 365 PLFDGIFCS 373

RESULT 15
US-10-062-809-3
Sequence 3, Application US/10062809
Publication No. US20030100106A1
GENERAL INFORMATION:
APPLICANT: Chang, Sandra P.
APPLICANT: Hashimoto, Anna
APPLICANT: Nishimura, Tani
TITLE OF INVENTION: BACULOVIRUS PRODUCED PLASMODIUM FALCIPARUM VACCINE
FILE REFERENCE: A-67984-1/RT/TAL/NBC
CURRENT APPLICATION NUMBER: US/10/062,809
CURRENT FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US 09/500,376
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: US 60/266,281
PRIOR FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 394
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-062-809-3

Query Match 56.1%; Score 1156.5; DB 9; Length 394;
Best Local Similarity 59.0%; Pred. No. 1e-66;
Matches 219; Conservative 54; Mismatches 73; Indels 25; Gaps 3;

QY 30 IDNLSKIEEYEVLYLKLACGYRSLLKQLENNVTFENVVYKDIINSRPNREKFNVL 89
DB 6 MNIIISGFENEYDVLYIKLACGYRSLLKQLENNVTFENVVYKDIINSRPNREKFNVL 65

QY 90 ESDLIPIKDLTSSNVVVDPIYKFLNKKRDKFLSSYNTIKDSIDPIDINPANDVLYGYKTL 149
DB 66 ESDLIQFKHISSENEYIIEDSFKLNSSEQKNTLKSSTYKIKESVENDIKFAQGISYERY 125
QY 150 SEKYKSDLSIKRYI-----NDKQGENEKYLPFLNIIETLYKTVN 189
DB 126 LAKYKDDLESIKKVIKEKEKFPSSPPTPPSPAKTDEQKKESKFLPFLNIIETLYKTVN 185
QY 190 HKIDLFVHLKAKVLYTEKSNVEYKIKELNYLKTIDKLADFKNKNNFVGIADLSTDY 249
DB 186 NKIDYLLIMKAKINDCNVEKDEAHYKTKLSDIAIDKIDLFKNTNDFEAIKRLIND 245
QY 250 NNNMLTKFLSTGMV--FENLAKTVLSNLLDGNLQGMNISOHCYKQCPONGSGCFRHL 307
DB 246 TKKDMGLKSLSTGLV--QNPNTIISKLFQDMNISOHCYKQCPONGSGCFRHL 302
QY 308 DERECKCLLNTKQEGDKCVENPPTCNENNGGCDADAKCTEEDSGSNKKTTCCTKPD 367
DB 303 DERECKCLLNTKQEGDKCVENPPTCNENNGGCDADATCTEEDSGSNKKTTCCTKPD 362
QY 368 SYPLFDGIFCS 378
DB 363 SYPLFDGIFCS 373

Search completed: June 11, 2003, 16:12:02
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: June 11, 2003, 16:02:30 ; Search time 40 Seconds
(without alignments)
922.890 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSWKCLFMAVLVTATLCT.....KPDSTPLFDGIFCSHHHHH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	91.1	400	2	A45545 major merozoite su
2	1879	91.1	1639	2	S05603 major merozoite su
3	1726.5	83.7	1631	1	SAZOK1 major merozoite su
4	1684.5	81.7	1640	2	A24594 probable major sur
5	1196.5	58.0	651	2	S47282 merozoite surface
6	1193	57.8	1726	1	SAZQGM major merozoite su
7	1186	57.5	1726	2	AA5948 major merozoite su
8	1174	56.9	1701	2	AA4498 major merozoite su
9	1158	56.1	1701	2	A26868 major merozoite su
10	743.5	36.0	1726	2	A39401 merozoite surface
11	742.5	36.0	1731	2	AA5604 major blood-stage
12	564.5	27.4	680	2	A28121 major merozoite su
13	564.5	27.4	1772	2	AA5532 major merozoite su
14	546.5	26.5	1785	2	A45546 major merozoite su
15	168.5	8.2	980	2	E71606 hypothetical prote
16	161	7.8	3394	2	T18501 hypothetical prote
17	152.5	7.4	1939	2	T18372 repeat organellar
18	151.5	7.3	481	2	T18465 hypothetical prote
19	150.5	7.3	1306	2	T28313 ORF MSV152 probabl
20	149.5	7.2	706	2	A57719 dihydropterolate sy
21	149	7.2	1714	2	E71609 Ser/Tth protein ki
22	148.5	7.2	2013	2	C71610 probable membrane
23	146.5	7.1	1619	2	T18499 hypothetical prote
24	146.5	7.1	2539	2	B71619 hypothetical prote
25	146.5	7.1	2708	2	T09079 probable chloroqui
26	146.5	7.1	2819	2	T09080 probable chloroqui
27	146	7.1	1127	2	T28317 ORF MSV156 hypotne
28	145.5	7.1	3844	2	T18402 asparagine/asparta
29	143.5	7.0	722	2	D71607 VPS45-like protein

30	143.5	7.0	1516	2	E71619 RAD2 endonuclease
31	143.5	7.0	1712	2	C71618 hypothetical prote
32	143.5	7.0	2500	2	G71609 hypothetical prote
33	141.5	6.9	662	2	C40862 heparin-binding gr
34	141.5	6.9	822	1	TYVHPC fibroblast growth
35	141.5	6.9	1005	2	A64465 hypothetical prote
36	141	6.8	1711	2	T18429 hypothetical prote
37	141	6.8	2401	2	T28676 rhoptry protein -
38	140	6.8	2910	2	T28156 DNA-directed RNA p
39	139.5	6.8	361	2	PN0020 fibroblast growth
40	139.5	6.8	822	2	I49289 fibroblast growth
41	139.5	6.8	1121	2	F71613 hypothetical prote
42	139.5	6.8	1979	2	C71622 hypothetical prote
43	139	6.7	2269	2	T28677 rhoptry protein -
44	138.5	6.7	707	2	T28418 ORF MSV257 leucine
45	138.5	6.7	822	1	TYVHPC fibroblast growth

ALIGNMENTS

RESULT 1

A45545 major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (fragments
C:Species: Plasmodium falciparum
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jun-2000
C:Accession: A45545

R.Blackman, M.J.; Ling, I.T.; Nicholas, S.C.; Holder, A.A.

Mol. Biochem. Parasitol. 49, 29-33, 1991

A:Title: sequence extracted from NCBI backbone (NCBIN:77612, NCBIPI:77621)

A:Reference number: A45545; MUID:92131048; PMID:1775158

A:Accession: A45545

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <BL>

A>Note: sequence extracted from NCBI backbone (NCBIN:77612, NCBIPI:77621)

A:Superfamily: major merozoite surface antigen

C:Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Query Match	91.1%	Score 1879;	DB 2;	Length 400;
Best Local Similarity	99.4%	Pred. No. 1.7e-92;		
Matches 355;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	22	GAATPSPYDNLISKEENEYEVLYLPLAGYRSIKOLENNVMTFNVNVKIDLSRPNK	81	
Db	23	GAATPSPYDNLISKEENEYEVLYLPLAGYRSIKOLENNVMTFNVNVKIDLSRPNK	82	
QY	82	RENFKNVLESDDIPYKDLTSSNVVVKDPYKFLNKKRDKFLSSVYIKDSIDTDFINFAND	141	
Db	83	RENFKNVLESDDIPYKDLTSSNVVVKDPYKFLNKKRDKFLSSVYIKDSIDTDFINFAND	142	
QY	142	VIGYKKIISSEKYSKSLDSIKRYINDKOGENEYVLFNNIETLYKTVNHNKIDLFYIHLEA	201	
Db	143	VIGYKKIISSEKYSKSLDSIKRYINDKOGENEYVLFNNIETLYKTVNHNKIDLFYIHLEA	202	
QY	202	KVLNTYKSNVEVRIKELNYLKTIDOKLADFKKNNNFVGIADLSTVDNHNHLLTKFPLST	261	
Db	203	KVLNTYKSNVEVRIKELNYLKTIDOKLADFKKNNNFVGIADLSTVDNHNHLLTKFPLST	262	
QY	262	GWFFENLAKTVLSNLDGNLQGLNISOHQCYKCKCPONSGCFRHLDERECCCLLNKQ	321	
Db	263	GWFFENLAKTVLSNLDGNLQGLNISOHQCYKCKCPONSGCFRHLDERECCCLLNKQ	322	
QY	322	BDKCVENPNCNNNGCGDADACTEDSSNGKKTCTCTKPDSTPLPFGIFCS	378	
Db	323	BDKCVENPNCNNNGCGDADACTEDSSNGKKTCTCTKPDSTPLPFGIFCS	379	
RESULT 2				
S05603				
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)				
N:Alternate names: gp195 surface antigen				
;Species: Plasmodium falciparum				

RESULT 2

S05603 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
N:Alternate names: gp195 surface antigen
C:Species: Plasmodium falciparum

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jun-2000
C/Accession: S05603; S04850
R/Myler, P.J.
Submitted to the EMBL Data Library, April 1989
A/Reference number: S05603
A/Accession: S05603
A/Molecule type: mRNA
A/Residues: 1-1639 <MYL>
A/Cross-references: EMBL:X15063; NID:99896; PID:CAA33163.1; PID:99897
R/Myler, P.J.
Nucleic Acids Res. 17, 5401, 1989
A/Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plasmodium falciparum
A/Reference number: S04850; MUID:89345116; PMID:2668887
A/Accession: S04850
A/Molecule type: mRNA
A/Residues: 1504-1639 <MYL2>
A/Cross-references: EMBL:X15063
C/Superfamily: major merozoite surface antigen
C/Keywords: glycoprotein; merozoite; surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 91.1%; Score 1879; DB 2; Length 1639;
Best Local Similarity 99.4%; Pred. No. 8.4e-92;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 81
DB 1262 GEAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 1321
QY 82 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 141
DB 1322 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 1381
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 201
DB 1382 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 1441
QY 202 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 261
DB 1442 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 1501
QY 262 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 321
DB 1502 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 1561
QY 322 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 378
DB 1562 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 1618

RESULT 3

SAZK1
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (str. SAZK1)
C/Species: Plasmodium falciparum
C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jun-2000
C/Accession: A25120
R/McKay, M.; Goman, M.; Bone, N.; Hyde, J.E.; Scaife, J.; Certa, V.; Stunnenberg, H.; E. MBO J. 4, 3833-3829, 1985
A/Title: Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum
A/Reference number: A91030; MUID:66136024; PMID:3004972
A/Accession: A25120
A/Molecule type: DNA
A/Residues: 1-1631 <MAC>
C/Comment: The merozoite stages of different strains have strain-specific surface antigen C. The merozoite has three stages: sporozoite, merozoite, and gametocyte. The merozoite stage is the major merozoite surface antigen.
C/Superfamily: major merozoite surface antigen
C/Keywords: glycoprotein; merozoite; surface antigen; tandem repeat; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1631/Product: major merozoite surface antigen #status predicted <MAT>
F:67-84/Region: 3-residue repeats (S-G-T/P)
F:1614-1631/Domain: membrane anchor #status predicted <MBN>
F:97,259,755,835,911,955,1049,1156,1165,1436,1563/Binding site: carbohydrate (Asn) (Asn)

Query Match 83.7%; Score 1726.5; DB 1; Length 1631;
Best Local Similarity 92.7%; Pred. No. 9.7e-84;
Matches 332; Conservative 6; Mismatches 19; Indels 1; Gaps 1;

QY 22 GAAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 81
DB 1253 GEAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 1312
QY 82 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 141
DB 1313 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 1372
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 201
DB 1373 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 1432
QY 202 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 261
DB 1433 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 1492
QY 262 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 320
DB 1493 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 1552
QY 321 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 378
DB 1553 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 1610

RESULT 4

A24594
probable major surface antigen (83k, 19k, 42k) precursor - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 09-Jun-2000
C/Accession: A24594
R/Holmes, A.A.; Lockyer, M.J.; Odink, K.G.; Sandhu, J.S.; Riveros-Moreno, V.; Nicholl Nature 317, 270-273, 1985
A/Title: Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum
A/Reference number: A24594; MUID:86014355; PMID:2995820
A/Accession: A24594
A/Molecule type: DNA
A/Residues: 1-1640 <HOL>
C/Superfamily: major merozoite surface antigen
C/Keywords: surface antigen

Query Match 81.7%; Score 1684.5; DB 2; Length 1640;
Best Local Similarity 91.3%; Pred. No. 1.6e-81;
Matches 327; Conservative 6; Mismatches 24; Indels 1; Gaps 1;

QY 22 GAAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 81
DB 1262 GEAVTPSVINDILISKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFVNVVKDILNSRPNK 1321
QY 82 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 141
DB 1322 RENNKNVLESPLIPYKDLTSSNYYKDPYFLNKKDKFLSSYNTKDSIDTDINPAND 1381
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 201
DB 1382 VLGYYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLNITETLYKYNHKTIDLFVHLEA 1441
QY 202 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 261
DB 1442 KVLNYYEKSNEVEKIKELNYLKTIDQKLADFFKNNFVGIADISTDYNNHNLTKFLST 1501
QY 262 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 320
DB 1502 GMVFNELAKYVLSLMDGNL-OGMLNISOHCYKQCPQNSGCFRHLDREBECKCLNLYK 1561
QY 321 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 378
DB 1562 EGDKCVENPPTCENNNGGCDADAKCTEEDSGSGNGKKTCECTKPDSPYLPFDGIFCS 1619

RESULT 5
S47282
merozoite surface antigen 1 - malaria parasite (Plasmodium falciparum) (strain RO-71) (4
C/Species: Plasmodium falciparum
A/Variety: strain RO-71
C/Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jun-2000
C/Accession: S47282
R/Tolle, R.; Bujard, H.; Cooper, J.A.
submitted to the EMBL Data Library, July 1994
A/Description: Plasmodium falciparum: recombination within the C-terminal region of mer
A/Reference number: S47282
A/Accession: S47282
A/Molecule type: DNA
A/Residues: 1-651 <TOL>
A/Cross-references: EMBL:235329; NID:9535257; PID:CAA84558.1; PID:9535258
A/Experimental source: strain RO-71
C/Superfamily: major merozoite surface antigen
C/Keywords: glycoprotein; merozoite; surface antigen

Query Match 58.0%; Score 1196.5; DB 2; Length 651;
Best Local Similarity 59.8%; Pred. No. 3.6e-56;
Matches 225; Conservative 56; Mismatches 74; Indels 21; Gaps 3;
QY 22 GAAVTPSVIDNLSIKENYEVLTKPLAGVRSLSKQLENNVMTFNVVVKILNSRFNK 81
DB 257 GEASVTV-KDNILSGFENEVDYIYKPLAGVRSLSKQLEKNIETFNLMNDILNSRLK 315
QY 82 REMFKNVLESDDLPIYDNLSSNVVADPYKFLKKEKRDKLSYNTIKSIDTDINPAND 141
DB 316 RKFFLDVLESDDLQFPHISSNEYIIEBSFKILNSEQKNTLLSKYKIKESVENDIKFAOE 375
QY 142 VLGYKYLSEKYSKSDLSIKKYI-----NDKOGENEKYLPLNNT 182
DB 376 GISYEVKVLAKYDDLESIKYVKEKEKPPSPPTTPSPAKTDEQKESKPLPLNTIE 435
QY 183 TLKYVNHKIDLFVILHLEKAVLNTYKESNVEYKIKELNYLTIDKIDAFKNNFVG 242
DB 436 TLVNNLVNKHIDYILNKKAKINDCNVEKDEAHVYKIKLSDLAKAIDKIDLFKNNHFEAL 495
QY 243 ADLSTVDYNNNNLITFLSTGWFENLAKYVLSNLDGNGMLNISQHCYKQCPONS 302
DB 496 KKLINDYTKKMDGLSTGLV-QNFPNTIISKLISGKFDMDLNTISQHCYKQCPONS 554
QY 303 CERHLDERECCCLNLYKQEGDKCVENPPTCENNGGCDADAKCTEEDSGSNGKRITCE 362
DB 555 CEFHLDERECCCLNLYKQEGDKCVENPPTCENNGGCDADAKCTEEDSGSNGKRITCE 614
QY 363 CTRKPDSTPLFDGIFCS 378
DB 615 CTRKPDSTPLFDGIFCS 630

RESULT 6
S40GM
Major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
N/Alternate names: 195K glycoprotein
C/Species: Plasmodium falciparum
C/Date: 30-Sep-1987 #sequence_revision 31-Mar-1991 #text_change 09-Jun-2000
C/Accession: A23386; S06361
R/Weber, J.L.; Leininger, W.M.; Lyon, J.A.
Nucleic Acids Res. 14, 3311-3323, 1986
A/Title: Variation in the gene encoding a major merozoite surface antigen of the human m
A/Reference number: A23386; M0ID:86205236; PMID:3517809
A/Accession: A23386
A/Molecule type: DNA
A/Residues: 1-1104 <WEB1>
A/Cross-references: EMBL:X03831
R/Weber, J.L.; Sim, B.K.L.; Lyon, J.A.; Wolff, R.
Nucleic Acids Res. 16, 1206, 1988
A/Title: Merozoite surface protein sequence from the Camp strain of the human malaria pa
A/Reference number: S06361; M0ID:86143999; PMID:3278296

A/Accession: S06361
A/Molecule type: DNA
A/Residues: 1104-1726 <WEB2>
A/Cross-references: EMBL:X03831
C/Comment: The merozoite stages of different strains have strain-specific surface ant
C/Comment: P. falciparum has three stages: sporozoite, merozoite, and gametocyte. The
C/Superfamily: major merozoite surface antigen
C/Keywords: glycoprotein; merozoite; surface antigen; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1726/Product: major merozoite surface antigen #status predicted <MAT>
F:67-87,91,96,100-105,109-120/Region: 3-residue repeats (S-G-T)
F:157-765/Region: 3-residue repeats (T-E-E)
F:133,272,501,567,638,827,839,924,944,990,1016,1114,1221,1613,1658/Binding site: carb

Query Match 57.8%; Score 1193; DB 1; Length 1726;
Best Local Similarity 59.4%; Pred. No. 1.7e-55;
Matches 224; Conservative 57; Mismatches 74; Indels 22; Gaps 3;
QY 22 GAAVTPSVIDNLSIKENYEVLTKPLAGVRSLSKQLENNVMTFNVVVKILNSRFNK 81
DB 1331 GEASVTV-KDNILSGFENEVDYIYKPLAGVRSLSKQLEKNIETFNLMNDILNSRLK 1389
QY 82 REMFKNVLESDDLPIYDNLSSNVVADPYKFLKKEKRDKLSYNTIKSIDTDINPAND 141
DB 1390 RKFFLDVLESDDLQFPHISSNEYIIEBSFKILNSEQKNTLLSKYKIKESVENDIKFAOE 1449
QY 142 VLGYKYLSEKYSKSDLSIKKYI-----NDKOGENEKYLPLNNT 181
DB 1450 GISYEVKVLAKYDDLESIKYVKEKEKPPSPPTTPSPAKTDEQKESKPLPLNTIE 1509
QY 182 ETLKYVNHKIDLFVILHLEKAVLNTYKESNVEYKIKELNYLTIDKIDAFKNNFVG 241
DB 1510 ETLVNNLVNKHIDYILNKKAKINDCNVEKDEAHVYKIKLSDLAKAIDKIDLFKNNHFEAL 1569
QY 242 IADLSTVDYNNNNLITFLSTGWFENLAKYVLSNLDGNGMLNISQHCYKQCPONS 301
DB 1570 IKKLINDYTKKMDGLSTGLV-QNFPNTIISKLISGKFDMDLNTISQHCYKQCPONS 1628
QY 302 CERHLDERECCCLNLYKQEGDKCVENPPTCENNGGCDADAKCTEEDSGSNGKRITC 361
DB 1629 CEFHLDERECCCLNLYKQEGDKCVENPPTCENNGGCDADAKCTEEDSGSNGKRITC 1688
QY 362 ECTKPDSTPLFDGIFCS 378
DB 1689 ECTKPDSTPLFDGIFCS 1705

RESULT 7
A45948
Major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C/Accession: A45948
R/Chang, S.P.; Kramer, K.J.; Yamaga, K.M.; Kato, A.; Case, S.E.; Siddiqui, W.A.
Exp. Parasitol. 67, 1-11, 1988
A/Title: Plasmodium falciparum: gene structure and hydrophathy profile of the major me
A/Reference number: A45948; M0ID:89005525; PMID:3049134
A/Accession: A45948
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1726 <CHN>
A/Cross-references: GB:M37213
C/Superfamily: major merozoite surface antigen
C/Keywords: surface antigen
Query Match 57.5%; Score 1186; DB 2; Length 1726;
Best Local Similarity 59.2%; Pred. No. 4e-55;
Matches 223; Conservative 57; Mismatches 75; Indels 22; Gaps 3;
QY 22 GAAVTPSVIDNLSIKENYEVLTKPLAGVRSLSKQLENNVMTFNVVVKILNSRFNK 81
DB 1331 GEASVTV-KDNILSGFENEVDYIYKPLAGVRSLSKQLEKNIETFNLMNDILNSRLK 1389

[illegible]

RESULT 8

A54498

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (Is

C:Species: Plasmodium falciparum

C:Date: 28-Oct-1994 #sequence-revision 28-Oct-1994 #text-change 09-Jun-2000

C:Accession: A54498

R:Peterson, M.G.; Coppel, R.L.; McIntyre, P.; Langford, C.J.; Woodrow, G.; Brown, G.V.;

Mol. Biochem. Parasitol. 27, 291-302, 1988

A:Title: Variation in the precursor to the major merozoite surface antigens of Plasmodi

A:Reference number: A54498; MUID:88142999; PMID:2449612

A:Accession: A54498

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1701 <P>

A:Cross-references: GB:M19143; NID:g160412; PIDN:AAA29653.1; PID:g160413

C:Superfamily: major merozoite surface antigen

C:Keywords: surface antigen

Query Match 56.9%; Score 1174; DB 2; Length 1701;

Best Local Similarity 58.6%; Pred. No. 1.7e-54;

Matches 221; Conservative 58; Mismatches 76; Indels 22; Gaps 3;

QY 22 GAATPSPVINDITSEINEYEVLYLKPPLAGYRSLKQLENNVATPNNVNYDILNSRPNK 81

Db 1306 GEALISVT-MDDNITLSEGFERNVDVILKPLAGYRSRLKQLENNITPNNLNDLILNSRKK 1364

QY 82 RENNKNVLESDLIYKPKDLTSSNVYKDPYKLNKKRDKRFLTSYNYIKDSIDTQINPAND 141

Db 1365 RKYFLDVLESDLMQFKHISSEYIETEDSFKLNSBQKNTLLKSYKIKESYENDIKRAQE 1424

QY 142 VLGYKILSEKYSKDLDSIKRYI-----NDKQGENEKYLPFLNNI 181

Db 1425 GISYIEKYLAKKDDLESIKRYIKKEEKEKFPSPSPPTPPSPAKDQEKESKFLPLTNI 1484

QY 182 ETLAKTYNNHKIDFLVHLEAKVNLVITYEKSVEYKIKELNLTAKTIQDKLADFKKNNFVG 241

Db 1485 ETLVNNLVNKKIDVILNFKALINDCYNVEKQEAHYKIKTLDLKAIDKIDIFKNTNDFEA 1544

QY 242 IADISTDYNNHNNLTFRFLSTGMVFENLAKTVLSLNLGNLOGMLNISOHOCVKKCPONS 301

Db 1545 IKTLLINDTKKMDLTKGLISTGLV-QNRPNTIISLILGKRFQDMNISOHOCVKKQCPENS 1603

QY 302 GCFPHLDREBECCKILNYKQEGDKCVENPPTCENNNGGCDADAKTEEDSGSNGKRTTC 361

Db 1604 GCFPHLDREBECCKLLNYKQEGDKCVENPPTCENNNGGCDADATCTIEEDSGSSRKRTTC 1663

QY 362 ECTKPDSTPLPFDGIFCS 378

|||||

```

Db          1664 ECTKPDGYPLEDGIIFS 1680

RESULT 9
A:26868
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jun-2000
C:Accession: A26868
R:Tanabe, K.; Mackay, M.; Goman, M.; Scaife, J.G.
J. Mol. Biol. 195, 273-287, 1987
A:Title: Allelic dimorphism in a surface antigen gene of the malaria parasite Plasmod
A:Reference number: A26868; MUID:88011243; PMID:3079521
A:Accession: A26868
A:Molecule type: DNA
A:Residues: 1-1701 <TRAN>
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1701/Product: major merozoite surface antigen #status predicted <MAT>

```

```

QY      22  AAATPSPVIONILSKIESEYEVLTAKPLAGYRSLSKKOLENNVTFPVNVYDLINSPNK 81
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1306 GEALSTV- MDNILSGFENEYDVLTKPLAGYRSLSKKOLENNITFPLNMDILNLSRLK 1364
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      82  RENEKNVLESDLIPKDLTSSNVYVKPKYKLNKKRDKFLSSYNYIKDSIDTDINPAMD 141
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1365 RKYPLVDLESDLMQFKHSSNEIYITDSFKLNSEQNNILKSKYIKIESVENDIKFAQE 1424
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      142  VLGYYKTLSEKYEKDDSIKKYI-----NDKGENEKYLPFLNTI 181
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1425 GISYEEYVLAAYKDKDLESIKKVIKEEKPPSPPTPPSPAKYDEKRSKFLPFLNTI 1484
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      182  ETLKYTVNHKIDLEFVILHEAKVLYEYKSNVEYKEIKNLTIOQLADFKKNNPVG 241
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1485 ETLVNNLVNKKIDVDYLINLKKAKINCQVNEKDAHAKITKLSDLKAIIDKIDLFKNTDPEA 1544
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      242  IADISTDYNNHNLITKFLSTGMY--FENLAKTVLSNLLDGNLQGLNISOHQYKCKOCP 299
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1545 IKKILINDYTKMDMGKLTSTGLVOIFPN---TIISKLTIEGFQDMLNISOHQCYKCKOPE 1601
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      300  NSGCFRHLDEREECKLLNTYKQEGDKCVENPNPTCNNGSGCDADACTEEDSGSSNGKI 359
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1602 NSGCFRHLDEREECKLLNTYKQEGDKCEBNPNPTCNENNGSCDADACTEEDSGSSNRKI 1661
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      360  TCCECTKPDSPYPLFDGIFCS 378
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1662 TCCECTKPDSPYPLFDGIFCS 1680
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
A39401
merozoite surface antigen 1 precursor - Plasmodium vivax
C:Species: Plasmodium vivax
C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Nov-1999
C:Accession: A39401
C:Idel Portillo, H.A.; Longacre, S.; Khouri, E.; David, P.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 4030-4034, 1991
A:Title: Primary structure of the merozoite surface antigen 1 of Plasmodium vivax rev
A:Reference number: A39401; MUID:91219506; PMID:2023952
A:Accession: A39401
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1726 <DEL>
A:Cross-references: GB:M60807
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

Query Match      36.0%; Score 743.5; DB 2; Length 1726;

```

[illegible]

```

RESULT 11
A:Accession: A45604
major blood-stage surface antigen Pv200 - Plasmodium vivax
C:Species: Plasmodium vivax
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C:Accession: A45604
R:Gibson, H.L.; Tucker, J.E.; Kaslow, D.C.; Kretzli, A.V.; Collins, W.E.; Klefer, M.C.;
Mol. Biochem. Parasitol. 50, 325-333, 1992
A:Title: Structure and expression of the gene for Pv200, a major blood-stage surface antigen
A:Reference number: A45604; MUID:92158013; PMID:1371329
A:Accession: A45604
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1751 <GIB>
A:Cross-references: GB:M75674; NID:g160608; PID:g457336
A>Note: sequence extracted from NCBI backbone (NCBIN:83591, NCBIPI:83592)
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

```

```

Query Match          36.0%: Score 742.5; DB 2; Length 1751;
Best Local Similarity 39.5%: Pred. No. 1.1e-31;
Matches 146; Conservative 71; Mismatches 112; Indels 41; Gaps 4;

QY  NEYEVLYLKPAGYVRSLSKQLENNVMVFNNVNDIINSRPNKRENFKNVLESDLIPYKD 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  SDYDVYVLKPLAGHYKTKIKQLEHNVNAFNNTINIDMLDSRLKRNRYELEVLNSDLNPKKY 1430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY  LTSNNVYVKKDPYKFLNKKERKDPKFLSSNYIKDSITDITDINFANDVLGYKIIISERYKSDLD 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  SSSGEYIHKDPYKLLIDELKKRKLIGSYKYGASIDMDLATINDGVTYNNKMGLEYLTKRIID 1490
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY  SIR-----KYIND-----KOGENEYLPFLNNIETLYT 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  GVKEIKKVEDIDIKQDEELKKLGNVSQDSKKKEFLAKKKELEKTYLPFLNSIQEWEESL 1556
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY  NKKIDLEVIHLEAFVNTYTERKSNEYVAKIRKELANTLYTIQDLADFKKNNNFVGIADLSTD 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  VSKVNTYTDNLKVIYNNQLEKKEAEIYTVKLDQYNNKDEKLELEYKK-----SEKNE 1603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY  YNNHNNLLTKFLSTGCMVFENLAKTVLSLIDSGNLOGLMINSIHOCHYKQCCPONSGCFRILD 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	1604	VKSSGJLEKMKSKLIRKNESEKILLSQLANVOQLUMSSBHTCIDTVPNMAOYRIID	1663
QY	309	ERECKCILLANKOGSDCVENPMPNCCNNNGCCADACCTEEDSGSNKRTITCETKXDS	368
Dd	1664	GTEWRCLITIKTEBEGKCVAPASVNTCKDNGSGCAPEAECKMTDS---NKIYCKCTKEGS	1719
QY	369	YLPFDGIFCS	378
Dd	1720	EPLEGEVPCS	1729

RESULT 12
A:28121
major merozoite surface antigen - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 26-Aug-1999
C:Accession: A28121
R:Burns Jr., J.M.; Daly, T.M.; Valdivya, A.B.; Long, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 602-606, 1988
A:Title: The 3' portion of the gene for a Plasmodium yoelii merozoite surface antigen
A:Reference number: A28121; MUID:88124889; PMID:2448778
A:Accession: A28121
A:Molecule type: DNA
A:Residues: 1-680

A:Cross-references: GB:J03612; NID:g160678; PID:g160679
A:Experimental source: strain 17XL
A:Note: the authors translated the codon GTA for residue 429 as leu
C:Superfamily: major merozoite surface antigen
C:Keywords: surface antigen

```

Query Match      27.4%; Score 564.5; DB 2; Length 680;
Best Local Similarity 34.6%; Pred. No. 1e-22;
Matches 126; Conservative 75; Mismatches 133; Indels 30; Gaps 9;

QY      32 NILSKINNEYEVLYKPLAGYYSKSKQLENNVATFENVANVKDILNSKRFENFKVYLES 91
Db      309 DILSEFNEESLYVYTKRLSGTYSKSLKRMHREESTIKEDNTGNNNSQRNDLFVLSH 368
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      92 DLIPYKDLTSNYYVVKQPYKFLNKKRDKFLSYNYIKDSIDIDINFAVNLGYYKILSE 151
Db      369 ELDFKRLSTNKKYIYRNPYQLDDNDKDKQIVNKKYATKGINIEDIETTGIKFENKME 428
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      152 KYKSDLDLSIKRYI-----ND-KQGENEYIPLNNIELLYKTVNKKIDLFYIHLKARY 203
Db      429 VYNTQLAAVEQIITIAETNNDINKKEKKYIPILEDLKGLEYEQAAEYSEBELNRL 468
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      204 LNTYKESNVEVKIKELNLYTKTIQDKLADF---KKNNNVGIGADLSIDYNNHNLTKFL 259
Db      489 DNYENKAEKEEILTKNKEYIQIDEKIDFVEVHAENKHHASIA-----LNNLN 537
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      260 STGVAFNLLKATVYLSNLLDNLQMG--LNTS-OHQCV-KKQCPNNSGCFPHLDREERCCK 315
Db      538 KSGLVGSGESKRIILAKRL--NMDSMDLLGYDPKRVICVDITDIRKNACGFFDDNGTEBKC 555
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      316 LNTATQ--EGDKCVENPPTCNENNNGCGDADACTCEDDSGNSGKKITCECTKRPDSYPLDG 374
Db      596 LLGKKKEGNTCVENNPCTDINNCGCDPTASCONAESTENSKKIITCCKEPTPNAYEG 655
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      375 IFCS 378
Db      656 VFCS 659
      :|::|

RESULT 13
A45532
major merozoite surface antigen precursor - Plasmodium yoelii1
C/Species: Plasmodium yoelii1
C/Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-Aug-1999
C/Accession: A45532; A45531
R:Lewis, A.P.
Mol. Biochem. Parasitol. 36, 271-282, 1989
A/Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surfa
C/Reference number: A45532; MUID:90014981; PMID:2797063

```

RESULT 13
A45532
major merozoite surface antigen precursor - Plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 03-Jun-1993 #sequence-revision 03-Jun-1993 #text-change 26-Aug-1999
C:Accession: A45532; A45531
R:Lewis, A.P.
Mol. Biochem. Parasitol. 36, 271-282, 1989
A:Title: Cloning and analysis of the gene encoding the 230-kilodalton merozoite surfa
#:Reference number: A45532; MUID:90014981; PMID:2797063

A:Accession: A45532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1772 <LEM>
 A:Cross-references: GB:J04668; NID:9160492; PID:9160493
 R:Daly, T.M.; Burns Jr., J.M.; Long, C.A.
 Mol. Biochem. Parasitol. 36, 283-285, 1989
 A:Title: Precursor to the major merozoite surface antigen of Plasmodium yoelii: cloning
 A:Reference number: A45531; MUID:90014982; PMID:2797064
 A:Accession: A45531
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 454-1094 <DAL>
 A:Cross-references: GB:J03975; NID:9160081; PID:9160082
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 27.4%; Score 564.5; DB 2; Length 1772;
 Best Local Similarity 34.6%; Pred. No. 3e-22;
 Matches 126; Conservative 75; Mismatches 133; Indels 30; Gaps 9;

QY 32 NLSIKENYEVLKPLAGYRSLSKQLENNMTFENVVNDILNSRFRKREKRVLES 91
 DB 1401 DILSEFPNESLYVTKRLSGYTKSLKKHMLREFSTIKEDMTNGLNNSQKNDLEVLSH 1460
 QY 92 DLIYKDLTSSNYVVDPKFLNKEKDKFLSTNYIKSDITDDINRANVLYGKILSE 151
 DB 1461 ELDFKDLSTKRYIRNPYQLDDNDKDKQIVNLLKATKGINEDIETTDGKFEFNKVE 1520
 QY 152 KYKSDLSIKRYI-----ND-KQGENEKYLPFLNNIEPLTKVNNHIDLFVHLKAVY 203
 DB 1521 LYNQLAVKQKQITAEFTMDTKKEKKKIPILEDKGLYEGVGAEEYSEELQRL 1580
 QY 204 LNTYKESNVEVRIKELNYIKTIQDKLADF---KKNNFVGADLSTDYNNHLLKFL 259
 DB 1581 DNYKNEKAEFELTKNLEKYIQIDELDEFVEHAENKHIASIA-----LNNLN 1629
 QY 260 STGVFENLAKTVSNLLDGLQGM--LNIS-QHQC-V-KKCCPONSQCFRLDERECKC 315
 DB 1630 KSGLVGSGESKILLAKML--NMDQMDLLGVDPKHVCVDTFPIPNAGCFRDNQTEWRC 1687
 QY 316 LANKQ-EGDKCVENPPTCENNNGCDADACTEEDSGSKITCECTPDSYPLDFD 374
 DB 1688 LIGKKEGNTCVENNPPTCDINNGGCDPTASCONAESTENSKITICTCKEPTNAYEG 1747
 QY 375 IFCS 378
 DB 1748 VFCS 1751

RESULT 14

A45346
 major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
 C:Species: Plasmodium chabaudi chabaudi
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
 C:Accession: A45546
 R:Deleenslinder, W.; Hendrix, D.; Bendahan, N.; Hanegreets, J.; Briljs, L.; Hamers-Caste
 Mol. Biochem. Parasitol. 43, 231-244, 1990
 A:Title: Molecular cloning and sequence analysis of the gene encoding the major merozoite
 A:Reference number: A45546; MUID:91218805; PMID:2090945
 A:Accession: A45546
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1785
 A:Cross-references: GB:M34947; NID:9160597; PID:9160598
 C:Superfamily: major merozoite surface antigen
 C:Keywords: surface antigen

Query Match 26.5%; Score 546.5; DB 2; Length 1785;
 Best Local Similarity 31.5%; Pred. No. 2.7e-21;
 Matches 117; Conservative 77; Mismatches 153; Indels 25; Gaps 6;
 QY 21 AGAAVTSVDNLSIKENYEVLKPLAGYRSLSKQLENNMTFENVVNDILNSRFRN 80

DB 1404 AMASTDEEVNDIIDAFSENEEYITKSLGTYKSFKKHMLKERSMIKEDIMTGLNKYLE 1463
 QY 81 KRENKNTLESDDLIPYKDLTSSNYVVDPKFLNKEKDKFLSSYNYIKSDITDDINRAN 140
 DB 1464 KRNFLDVLSTELALFKQIDINTKVKNPQVLDDNDKQKQIMILKAIKQVTDIETAT 1533
 QY 141 DVLCGYKILSKRYSDLSIKRYI-----NDKQGENEKYLPFLNNIEPLTKVNNHIDLF 195
 DB 1524 DGIEFFNMIELYRQNALNAVEQJAAIGTEPTDAEKRYAIFEDLAGLYETIINGAEF 1583
 QY 196 VHLKAVLNTTYKESNVEVRIKELNYIKTIQDKLADF---KKNNFVGADLSTDYNNH 251
 DB 1584 SELQHKLENTYKIRKAGFDIIMALEYIRIDERLEDFVESAEKKNHIASIA----- 1635
 QY 252 NNLTKFLSTGVFENLAKTVSNLLDGLQGM--LNIS-QHQC-V-KKCCPONSQCFRLHDER 310
 DB 1636 ---LNNLNSGLVTEGESKILLAKMLNMDADLIGISNHYCISTPDDNNGCFRVDGT 1692
 QY 311 EECRCCLNTYKQ-EGDKCVENPPTCENNNGCDADACTEEDSGSN--GKRTCECTRP 366
 DB 1693 EEMKCLGFKKDDQGNRCVADAVCVNNNGGCDKNADCREVENTDPSKIIYCTQKEP 1752
 QY 367 DSYPLFDGIFCS 378
 DB 1753 PNNAVYAGVFCS 1764

RESULT 15

E71606
 hypothetical protein PFB0765w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 A:Accession: E71606
 R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Areavid, L.; Koonin, E.
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: E71606
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-980 <GAR>
 A:Cross-references: GB:AE001417; GB:AE001362; NID:93845271; PID:NACT1942.1; PID:9384
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0765w
 C:Superfamily: hypothetical protein MJ1322

Query Match 8.2%; Score 168.5; DB 2; Length 980;
 Best Local Similarity 24.5%; Pred. No. 0.13;
 Matches 88; Conservative 62; Mismatches 118; Indels 91; Gaps 21;

QY 34 LSKRIENYEVLKPLAGYRSLSKQLENNMTFENVVNDILNSRFRKREKRVLES 93
 DB 355 VNKIQNEL-----IKRELQNKCTIS-----KDLFECKKEDKIKN-LEDDL 394
 QY 94 IPRYDLTSSNYVVDPKFLNKEKDKR-FLSSYNYIKSDITDDINRANVLYGKILSE 151
 DB 395 LEKKKQJEN--LDELINIKKKMEDKMHNMENDLSNVE-ELANKINT--YERIVIE 448
 QY 152 KYKSDLSIKRYINDKO--GENEKYLPFLNNIEPLTKVNNHIDLFVHLKAVLNTTYE 209
 DB 449 -LNNELLVIRKKLDEDFLEKEEK---KKNIDVYKIRKEVEID--IKERENEDISLKN 501
 QY 210 KSNVEY-KIKELN-----YKTIQDKLADPKN-----NNFVGADLSTDYNNH 251
 DB 502 EONLHVLNKELEKELIILNKYDKELMILEIYQNNKIQOEKDLNKKI---KSMQDTH 557
 QY 252 NNLTKFLSTGVFENLAKTV-LSNLDGLQGM--LNIS-QHQC-V-KKCCPONSQCFRLHDER 310
 DB 558 KNOIEE-----MOENKKELEKRLKNVCDMLNQSILIKEN-----KHOEK 599

Thu Jun 12 08:47:26 2003

us-09-500-376-8.rpr

Page 7

```

QY      311 -EECCCLNKK-----QEDBCKVBNPNPTCNENNNGGCDADAKCTEEDSGSNKKIT 360
        ||||| |  |||:| | : | | |||
DB      600 VEEYNNLLKQDQELKNIIQFYDERIEIQNKEMEDIYVDCCEERUK---QAKINNKLT 654

```

Search completed: June 11, 2003, 16:06:39
Job time : 42 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:02:30 ; Search time 23 Seconds

(without alignments)
692,474 Million cell updates/sec

Title: US-09-500-376-8

Perfect score: 2063
Sequence: 1 MMSMKLLFMVAVLVATLCT.....KPDSTPLRDIQICSHHHHHH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	91.1	1630	1 MSPL_PLAIFK	P04932 plasmodium
2	1879	91.1	1639	1 MSPL_PLAIFV	P04933 plasmodium
3	1193	57.8	1726	1 MSPL_PLAIFC	P04934 plasmodium
4	1192	57.8	1726	1 MSPL_PLAIFP	P50495 plasmodium
5	1185.5	57.5	1682	1 MSPL_PLAIF3	P19598 plasmodium
6	1174	56.9	1701	1 MSPL_PLAIFR	P13819 plasmodium
7	1174	56.9	1701	1 MSPL_PLAIFM	P08569 plasmodium
8	564.5	27.4	1772	1 MSPL_PLAIFV	P13828 plasmodium
9	141.5	6.9	822	1 FGRI_HUMAN	P11362 homo sapien
10	141.5	6.9	1005	1 RAS0_METUA	O58718 methanococ
11	139.5	6.8	822	1 FGRI_MOUSE	P16092 mus musculu
12	138	6.7	1150	1 RRI1_YEAST	P40541 saccharomyc
13	137.5	6.7	1251	1 RBP2_PLAIVB	O00799 plasmodium
14	136.5	6.6	617	1 Y237_BUCAL	P57331 buchnera ap
15	135	6.5	537	1 ARP_PLAIFR	P04931 plasmodium
16	134.5	6.5	756	1 Y328_MYCGE	O49419 mycoplasma
17	133.5	6.5	455	1 EPIC_STAEP	P30196 staphylococ
18	132.5	6.4	377	1 Y704_METUA	O58115 methanococ
19	132.5	6.4	1076	1 RPOB_ASTLO	P27059 astasia lon
20	132.5	6.4	1113	1 Y140_MYCPN	P75033 mycoplasma
21	131.5	6.4	1030	1 Y018_MYCPN	P75093 mycoplasma
22	130	6.3	822	1 FGRI_RAT	O04589 rattus norv
23	129.5	6.3	971	1 Y228_BORBU	O51246 borrella bu
24	126	6.1	490	1 Y032_BORBU	O51063 borrella bu
25	125.5	6.1	1169	1 EX5B_BORBU	O51578 borrella bu
26	125.5	6.1	1169	1 SMC_METUA	O59037 methanococ
27	125.5	6.1	1228	1 ECN_METUA	O13201 homo sapien
28	125.5	6.1	2104	1 MYS3_SCHPO	O41457 schizosach
29	124.5	6.0	1108	1 CINH_YEAST	P27895 saccharomyc
30	124.5	6.0	1102	1 RPOP_AGART	P33399 agarticus bi
31	123.5	6.0	1790	1 US01_YEAST	P25386 saccharomyc
32	123.5	6.0	339	1 RMAR_CANCA	P21358 candida gla
33	123.5	6.0	1174	1 EX5B_BUCAL	P57529 buchnera ap

34	123.5	6.0	1956	1 ATX1_PLAIF	O04956 plasmodium
35	122.5	5.9	819	1 FGRI_CHICK	P21804 gallus gall
36	122	5.9	881	1 LHS1_YEAST	P36016 saccharomyc
37	121.5	5.9	478	1 DISR_AGRH	P30403 aglistrodon
38	121.5	5.9	656	1 V091_FOWPV	O72896 fowlpox vir
39	121.5	5.9	864	1 CHEA_BORBU	O44737 borrella bu
40	121.5	5.9	1162	1 BXEN_CLOBU	O06366 clostridium
41	121	5.9	1783	1 Y468_MYCGE	O49460 mycoplasma
42	120.5	5.8	1162	1 BXEN_CLOBO	P46082 clostridium
43	120.5	5.8	2198	1 Y132_CAEEL	P34367 caenorhabdi
44	120	5.8	540	1 YK66_YEAST	P36112 saccharomyc
45	120	5.8	571	1 V096_FOWPV	O935c4 fowlpox vir

ALIGNMENTS

```

RESULT 1
MSPL_PLAIFK STANDARD; PRT; 1630 AA.
ID MSPL_PLAIFK
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate KI / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stuenkelberg H., Bujard H.;
RT Polymorphism of the precursor for the major surface antigens of
RT Plasmodium falciparum merozoites: studies at the genetic level.*;
RL EMBO J. 4:3823-3829(1985).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RA Pan W., Tolle R., Bujard H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(Potential).
-1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sb.ch/announce/
or send an email to license@isb-sb.ch).
-----
PMB: X03371; CAA27070.1; -.
PIR: A25120; SAZOK1.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW Malaria; Merozoite; GPI-anchored; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1630 AA: 187289 MW; ADDECCGCEDA6322 CR664;

Query Match 91.1%; Score 1879; DB 1; Length 1630;
 Best Local Similarity 99.4%; Pred. No. 4.4e-95;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSVINDIISKIENEYVYLKPLAGYRSKQLNNVTFVNVKDIINSRPNK 81
 DB 1253 GEATPSVINDIISKIENEYVYLKPLAGYRSKQLNNVTFVNVKDIINSRPNK 1312
 QY 82 RENKNVLESDLIPYKLTSSNVYKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 141
 DB 1313 RENKNVLESDLIPYKLTSSNVYKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 1372
 QY 142 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYPPLNNIETLYKTVNHKIDLFVHLEA 201
 DB 1373 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYPPLNNIETLYKTVNHKIDLFVHLEA 1432
 QY 202 KVLNYYTEKSNVEYKIKELNLYKTIQDKLADFKNNNFVGIADISTDYNHNNLLTKRFLST 261
 DB 1433 KVLNYYTEKSNVEYKIKELNLYKTIQDKLADFKNNNFVGIADISTDYNHNNLLTKRFLST 1492
 QY 262 GMVENIAKTYVLSMLDGNLQGMNLISOHOCVVKOCPONSGCFRHHDERECCKCLLNYKQ 321
 DB 1493 GMVENIAKTYVLSMLDGNLQGMNLISOHOCVVKOCPONSGCFRHHDERECCKCLLNYKQ 1552
 QY 322 EGDCKVENPNTCTENNGGCDADAKCTEEDSGSNGKRITCECTRPDSYPLFDGIFCS 378
 DB 1553 EGDCKVENPNTCTENNGGCDADAKCTEEDSGSNGKRITCECTRPDSYPLFDGIFCS 1609

RESULT 2

MSPL_PLAFW STANDARD; PRT; 1639 AA.

AC P04933;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMSA) (P195).
 GN MSP-1.
 OS Plasmodium falciparum (isolate Wellcome).
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5848;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8601355; PubMed=2995820;
 RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
 RA Nicholson S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
 RA Freeman R.R.;
 RT "Primary structure of the precursor to the three major surface
 RT antigens of Plasmodium falciparum merozoites";
 RL Nature 317:270-273(1985).
 RN [2]
 RP REVISIONS.
 RA Holder A.A.;
 RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (potential).
 CC -1- PPM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 KDa, 42
 CC KDa AND 19 KDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X02919; CA26676.1; -
 DR PIR: A24594; A24594.
 DR InterPro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF 1.
 DR K1: Transmembrane; GPI-anchor.

FT CHAIN 1 19 POTENTIAL.
 FT SIGNAL 20 1639 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 768 768 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1639 AA: 187618 MW; 2C255B61C87F6E CR664;

Query Match 91.1%; Score 1879; DB 1; Length 1639;
 Best Local Similarity 99.4%; Pred. No. 4.4e-95;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSVINDIISKIENEYVYLKPLAGYRSKQLNNVTFVNVKDIINSRPNK 81
 DB 1262 GEATPSVINDIISKIENEYVYLKPLAGYRSKQLNNVTFVNVKDIINSRPNK 1321
 QY 82 RENKNVLESDLIPYKLTSSNVYKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 141
 DB 1322 RENKNVLESDLIPYKLTSSNVYKDPYKFLNKKRDKFLSSNYIKDSIDTDINPAND 1381
 QY 142 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYPPLNNIETLYKTVNHKIDLFVHLEA 201
 DB 1382 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYPPLNNIETLYKTVNHKIDLFVHLEA 1441
 QY 202 KVLNYYTEKSNVEYKIKELNLYKTIQDKLADFKNNNFVGIADISTDYNHNNLLTKRFLST 261
 DB 1442 KVLNYYTEKSNVEYKIKELNLYKTIQDKLADFKNNNFVGIADISTDYNHNNLLTKRFLST 1501
 QY 262 GMVENIAKTYVLSMLDGNLQGMNLISOHOCVVKOCPONSGCFRHHDERECCKCLLNYKQ 321
 DB 1502 GMVENIAKTYVLSMLDGNLQGMNLISOHOCVVKOCPONSGCFRHHDERECCKCLLNYKQ 1561
 QY 322 EGDCKVENPNTCTENNGGCDADAKCTEEDSGSNGKRITCECTRPDSYPLFDGIFCS 378
 DB 1562 EGDCKVENPNTCTENNGGCDADAKCTEEDSGSNGKRITCECTRPDSYPLFDGIFCS 1618

RESULT 3

MSPL_PLAFW STANDARD; PRT; 1726 AA.

AC P04934;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMSA) (P195).
 GN MSP-1.
 OS Plasmodium falciparum (isolate Camp / Malaysia).
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5835;
 RN [1]
 RP SEQUENCE OF 1-1103 FROM N.A.
 RX MEDLINE=86205236; PubMed=3517809;
 RA Weber J.L., Lehniger W.M., Lyon J.A.;

RT Variation in the gene encoding a major merozoite surface antigen of
 RT the human malaria parasite *Plasmodium falciparum*.
 RL Nucleic Acids Res. 14:3311-3323(1986).
 RN [2]
 RP SEQUENCE OF 1104-1726 FROM N.A.
 RX MEDLINE-88143999; PubMed-3278296;
 RA Weber J.L., Sim B.K.L., Lyon J.A., Wolff R.;
 RT "Merozoite surface protein sequence from the Camp strain of the human
 RT malaria parasite *Plasmodium falciparum*."
 RL Nucleic Acids Res. 16:1206-1206(1988).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X03831; CA27446.1; -
 DR PIR: A23386; SAZQM.
 DR InterPro: IPR000561; EGF-1-like.
 DR Pfam: PF00008; EGF_1.
 KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT CHAIN 1 19
 FT SIGNAL 1 19
 FT CARBOHYD 133 1726 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 272 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 501 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 567 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 827 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 924 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 944 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 990 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1016 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1114 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1221 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1613 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1726 1726 AA; 196197 MW; D08AD45FA352BCF3 CRC64;
 SQ SEQUENCE
 Query Match 57.8%; Score 1193; DB 1; Length 1726;
 Best Local Similarity 59.4%; Pred. No. 9.8e-58;
 Matches 224; Conservative 57; Mismatches 74; Indels 22; Gaps 3;
 QY 22 GAATPSVVDNLTISKENEYEVLYLKLPLAGYRSLSKKQLENNVMTFNVVVKILNSRFNK 81
 DB 1331 GEAISVT-MDNLTSGENEYDVYLYLKLPLAGYRSLSKKQLENNVMTFNVVVKILNSRLKK 1389
 QY 82 RENFNALVESDLIPYKDLTSSNVVVDPKFLNKKERDFLSYNYIKSDIDTDFINPAND 141
 DB 1390 RKYFLDVLSDLMQFHRISSENYIIDSFKLNSQKNTLLSKYIKESVENDIFAGE 1449
 QY 142 VLGYYKILSEKYSKSDLSIKTYL-----NDKQGENEYKLPPLNNI 181
 DB 1450 GISYKEVYLAQKDDLESIKYIKERKEFPSSPPTTPSPAKTDQKESKFLPLPLFNI 1509
 QY 182 ETLKYTNKRIDLEVHLFAKVLNTYKESNVEYKIKELNYTKITDIDKADFEKNPNFVG 241
 DB 1510 ETLNNLVNKKIDYLYLNLAKIKDCNVEDEARVKTITKSLDKAIDDKIDLFNNHDFEA 1569
 QY 242 IADLSDYNNHNLTKFLSTGWFFENLAVTLNLLDGNLQGNLTNSQHCYAKQCPONS 301
 DB 1570 IKKLINDPTKKMDLGLSTGLY-QNFPNTIISKLEIGKFDMLNTNSQHCYAKQCPONS 1628
 QY 302 GCFRHLDERBECKCLNTYKQEGDKCVENPNPCNENNGCDADAKTEEDSGSNGKKTTC 361
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1629 GCFRHLDERBECKCLNTYKQEGDKCVENPNPCNENNGCDADAKTEEDSGSNGKKTTC 1688
 QY 362 ECKPDPSTPLFDGIFCS 378
 DB 1689 ECKPDPSTPLFDGIFCS 1705
 RESULT 4
 ID MSP1.PLAFP STANDARD; PRT: 1726 AA.
 AC P50495;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
 DE (PMMA) (Gp195).
 GN MSP-1.
 OS Plasmodium falciparum (Isolate Palo Alto / Uganda).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89005525; PubMed-3049134;
 RA Chang S.P., Kramer K.J., Yamaga K.M., Kato A., Case S.E.,
 RA Siddiqui W.A.;
 RT "Plasmodium falciparum: gene structure and hydropathy profile of the
 RT major merozoite surface antigen (gp195) of the Uganda-Palo Alto
 RT isolate."
 RL Exp. Parasitol. 67:1-11(1988).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M37213; AAA29611.1; -
 DR InterPro: IPR000561; EGF-1-like.
 DR Pfam: PF00008; EGF_1.
 KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
 KW Transmembrane; GPI-anchor.
 FT CHAIN 1 19
 FT SIGNAL 1 19
 FT CARBOHYD 133 1726 MEROZOITE SURFACE PROTEIN 1.
 FT CARBOHYD 272 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 501 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 567 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 827 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 924 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 944 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 990 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1016 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1114 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1221 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1613 1726 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1726 1726 AA; 5859CEBFA2FPA026 CRC64;
 SQ SEQUENCE
 Query Match 57.8%; Score 1192; DB 1; Length 1726;
 Best Local Similarity 59.4%; Pred. No. 1.1e-57;
 Matches 224; Conservative 57; Mismatches 74; Indels 22; Gaps 3;
 QY 22 GAATPSVVDNLTISKENEYEVLYLKLPLAGYRSLSKKQLENNVMTFNVVVKILNSRFNK 81
 DB 1331 GEAISVT-MDNLTSGENEYDVYLYLKLPLAGYRSLSKKQLENNVMTFNVVVKILNSRLKK 1389

-1- PTM: MERZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42 kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MERZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: M19143; AAA29653.1; -
 CC PIR: A54498; A54498.
 CC InterPro: IPR000561; EGF-like.
 CC Pfam: PF00008; EGF; 1.
 CC Malaria: Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
 CC Transmembrane; GPI-anchor.
 CC SIGNAL 1 19
 CC CHAIN 20 1701 POTENTIAL.
 CC MERZOITE SURFACE PROTEIN 1.
 CC CARBOHD 110 110 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 239 239 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 470 470 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 536 536 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 607 607 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 802 802 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 899 899 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 919 919 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 965 965 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 991 991 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1089 1089 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1196 1196 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1588 1588 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC SEQUENCE 1701 AA; 193719 MW; 3920B75E73D38552 CRC64;

Query Match 56.9%; Score 1174; DB 1; Length 1701;
 Best Local Similarity 58.6%; Pred. No. 1e-56;
 Matches 221; Conservative 58; Mismatches 76; Indels 22; Gaps 3;

22 GAAVTPSYIDNLIKENEYEVLYLKPLAGYRSLSKOLENNVMTFNVVNDILNSRPNK 81
 1306 GEALISVT-MNDILSGFENEYVYILKPLAGYRSLSKOLENNITFNLNDILNSRLKK 1364

82 RENFNVLSEDLIPYKIDLTSSNNYVKKDFKFLNKKDKLTSSYNYKIDSDIDTDFNAND 141
 1365 RKFYLDVLESDIMQFKHISSENYIIEISFKLNSQKNTLTKSYKYESEVENDIRFAQE 1424

142 VIQYKILSEKYSKSDLSIKRYI-----NDKQENKYLPEFLNNT 181
 1425 GISYIEKVLAKYKDDLEIKRVKEKEKPESSPTTPPPSPAKTDEQKESKFLPEFLNTI 1484

182 ETLKRYVNHKIDLEVIHLEAKVLYTEKSNVEYKIKELANTKTQDILADFKKNNFVG 241
 1485 ETLNNNLVKNKIDYVLINKAKINOCNVEKDEAHVKITLSDLKAIIDKIDLEFKNNDEEA 1544

242 IADISTDYNNNNLLTKFLSTGMEFENTAKTYVLSNLDGNTLGMLNISQHCYKQCPONS 301
 1545 IRLKLNIDTKKMDKGLSTGLV-ONFENYIISKILIEKFPDMNINISQHCYKQCPENS 1603

302 GCFPHLDRECKCLNANTKQEGDCVENPPTCNENNGGCGADAKCTEEDSGSNKRTTC 361
 1604 GCFPHLDRECKCLNANTKQEGDCVENPPTCNENNGGCGADAKCTEEDSGSNKRTTC 1663

362 ECTKPDYSYPLFDGIFCS 378
 1664 ECTKPDYSYPLFDGIFCS 1680

RESULT 7
 MSP1_PLAFM STANDARD; PRT; 1701 AA.
 AC P08569;
 DT 01-AUG-1988 (Rel. 08, Created)

30-MAY-2000 (Rel. 39, Last sequence update)
 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Merzoite surface protein 1 precursor (merzoite surface antigens)
 GN (PMMsa) (P190).
 OS Plasmodium falciparum (isolate mad20 / Papua New Guinea).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID:70153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86011243; PubMed=3079521;
 RA Tanabe K., Mackay M., Goman M., Scaife J.G.;
 RT "Allelic dimorphism in a surface antigen gene of the malaria parasite
 RT Plasmodium falciparum.";
 RL J. Mol. Biol. 195;273-287(1987).
 [2]
 RP REVISIONS TO 1403; 1569 AND 1629.
 RA Tanabe K.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [3]
 CC SEQUENCE OF 1-115 FROM N.A.
 CC MEDLINE=86136024; PubMed=3004972;
 CC Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
 CC Stuenkelberg H., Bulard H.;
 CC "Polymorphism of the precursor for the major surface antigens of
 CC Plasmodium falciparum merzoites: studies at the genetic level.";
 CC EMBO J. 4:3823-3829(1985).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (Potential).
 CC -1- PTM: MERZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
 CC kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
 CC MERZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: X05624; CA29112.1; -
 CC PIR: A26868; A26868.
 CC InterPro: IPR000561; EGF-like.
 CC Pfam: PF00008; EGF; 1.
 CC Malaria: Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
 CC Transmembrane; GPI-anchor.
 CC SIGNAL 1 19
 CC CHAIN 20 1701 POTENTIAL.
 CC MERZOITE SURFACE PROTEIN 1.
 CC CARBOHD 110 110 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 239 239 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 470 470 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 536 536 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 607 607 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 802 802 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 899 899 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 919 919 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 965 965 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 991 991 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1089 1089 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1196 1196 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC CARBOHD 1588 1588 N-LINKED (GLCNAc . . .) (POTENTIAL).
 CC SEQUENCE 1701 AA; 193768 MW; 3FC2EC59AE96EA98 CRC64;

Query Match 56.9%; Score 1174; DB 1; Length 1701;
 Best Local Similarity 58.6%; Pred. No. 1e-56;
 Matches 221; Conservative 58; Mismatches 76; Indels 22; Gaps 3;

22 GAAVTPSYIDNLIKENEYEVLYLKPLAGYRSLSKOLENNVMTFNVVNDILNSRPNK 81
 1306 GEALISVT-MNDILSGFENEYVYILKPLAGYRSLSKOLENNITFNLNDILNSRLKK 1364

RP SEQUENCE FROM N.A.
 RC TISSUE-Neonatal brain stem;
 RX MEDLINE-90360977; PubMed-1697263;
 RA Dione C.A., Crumley G.R., Bellot F., Kaplow J.M., Searfoss G.,
 RA Ruta M., Burgess W.H., Jaye M., Schlessinger J.;
 RT "Cloning and expression of two distinct high-affinity receptors
 RT cross-reacting with acidic and basic fibroblast growth factors.";
 RL EMO J. 9:2685-2692(1990).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE-92282615; PubMed-1317750;
 RA Hattori Y., Odagiri H., Katoh O., Sakamoto H., Morita T.,
 RA Shimotohno K., Tobinai K., Sugimura T., Terada M.;
 RT "K-sam-related gene, N-sam, encodes fibroblast growth factor receptor
 RT and is expressed in T-lymphocytic tumors.";
 RL Cancer Res. 52:3367-3371(1992).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RX MEDLINE-91126480; PubMed-1846977;
 RA Hou J., Kan M., McKeenan K., McBride G., Adams P., McKeenan W.L.;
 RT "Fibroblast growth factor receptors from liver vary in three
 RT structural domains.";
 RL Science 251:665-668(1991).
 RN [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE-92118399; PubMed-1662973;
 RA Kleier M.C., Baird A., George-Nascimento C., Nguyen T., Mason O.B.,
 RA Boley L.J., Valenzuela P., Barr P.J.;
 RT "Molecular cloning of a human basic fibroblast growth factor receptor
 RT cDNA and expression of a biologically active extracellular domain in
 RT a baculovirus system.";
 RL Growth Factors 5:115-127(1991).
 RN [6]
 RN SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE-90290512; PubMed-2162671;
 RA Itton N., Terachi T., Ohta M., Seo M.K.;
 RT "The complete amino acid sequence of the shorter form of human basic
 RT fibroblast growth factor receptor deduced from its cDNA.";
 RL Biochem. Biophys. Res. Commun. 169:680-685(1990).
 RN [7]
 RN SEQUENCE OF 201-822 FROM N.A.
 RA Ruta M., Howk R., Ricca G., Drohan W., Zabelshansky M., Laureys G.,
 RA Barton D.E., Francke U., Schlessinger J., Glyvol D.;
 RT "A novel protein tyrosine kinase gene whose expression is modulated
 RT during endothelial cell differentiation.";
 RL Oncogene 3:9-15(1988).
 RN [8]
 RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE-90355989; PubMed-2167437;
 RA Johnson D.E., Lee P.L., Lu J., Williams L.T.;
 RT "Diverse forms of a receptor for acidic and basic fibroblast growth
 RT factors.";
 RL Mol. Cell. Biol. 10:4728-4736(1990).
 RN [9]
 RN ALTERNATIVE SPLICING.
 RX MEDLINE-91141499; PubMed-1847500;
 RA Gutkind S.J., Link D.C., Katamine S., Lacal P., Miki T., Ley T.J.,
 RA Robbins K.C.;
 RT "A novel c-fgr exon utilized in Epstein-Barr virus-infected B
 RT lymphocytes but not in normal monocytes.";
 RL Mol. Cell. Biol. 11:1500-1507(1991).
 RN [10]
 RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE-Lung;
 RX MEDLINE-91194400; PubMed-1650441;
 RA Bismann A., Ahn J.A., Graziani G., Tronick S.R., Ron D.;
 RT "Alternative splicing generates at least five different isoforms of
 RT the human basic-FGF receptor.";
 RL Oncogene 6:1195-1202(1991).
 RN [11]
 RN SEQUENCE FROM N.A.

RA Wennstroem S., Sandstroem C., Claesson-Welsh L.;
 RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RN MUTAGENESIS OF TYR-766.
 RX MEDLINE-92357144; PubMed-1379697;
 RA Peters K.G., Marie J., Wilson E., Ives H.E., Escobedo J.,
 RA del Rosario M., Mirza D., Williams L.T.;
 RT "Point mutation of an FGF receptor abolishes phosphatidylinositol
 RT turnover and Ca2+ flux but not mitogenesis.";
 RL Nature 358:678-681(1992).
 RN [13]
 RN MUTAGENESIS OF TYR-766.
 RX MEDLINE-92357145; PubMed-1379698;
 RA Mohammad M., Dione C.A., Li W., Lin N., Spivak T., Honneger A.M.,
 RA Jaye M., Schlessinger J.;
 RT "Point mutation in RGF receptor eliminates phosphatidylinositol
 RT hydrolysis without affecting mitogenesis.";
 RL Nature 358:681-684(1992).
 RN [14]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 464-762.
 RX MEDLINE-96361355; PubMed-8752212;
 RA Mohammad M., Schlessinger J., Hubbard S.R.;
 RT "Structure of the FGF receptor tyrosine kinase domain reveals a novel
 RT autoinhibitory mechanism.";
 RL Cell 86:577-587(1996).
 RN [15]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 464-762.
 RX MEDLINE-97284786; PubMed-9139660;
 RA Mohammad M., McMahon G., Sun L., Tang C., Hirth P., Yeh B.K.,
 RA Hubbard S.R., Schlessinger J.;
 RT "Structures of the tyrosine kinase domain of fibroblast growth factor
 RT receptor in complex with inhibitors.";
 RL Science 276:955-960(1997).
 RN [16]
 RN VARIANT PEPTIDE SYNDROME ARG-252.
 RX MEDLINE-95179173; PubMed-7874169;
 RA Muenke M., Scheil U., Hehr A., Robin N.H., Losken H.W., Schinzel A.,
 RA Pullen L.J., Rutland P., Reardon W., Malcolm S., Winter R.M.;
 RT "A common mutation in the fibroblast growth factor receptor 1 gene in
 RT Pfeiffer syndrome.";
 RL Nat. Genet. 8:269-274(1994).
 CC -1- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER
 CC FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (AFGF).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: MANY FORMS OF FGFR ARE PRODUCED BY
 CC ALTERNATIVE SPLICING. THE FORM SHOWN HERE IS KNOWN AS ALPHA-A1.
 CC -1- DISEASE: DEFECTS IN FGFR1 ARE ONE OF THE CAUSES OF PEIFFER
 CC SYNDROME (FS) (ALSO KNOWN AS ACROCEPHALOSYNDACTYL TYPE V; ACS5);
 CC CHARACTERIZED BY CRANIOSYNOSTOSIS (PREMATURE FUSION OF THE SKULL
 CC SUTURES) WITH DEVIATION AND ENLARGEMENT OF THE THUMB AND GREAT
 CC TOES, BRACHYMESEPHALANGY, WITH PHALANGEAL ANKYLOSIS AND A VARYING
 CC DEGREE OF SOFT TISSUE SYNDACTYL.
 CC -1- DISEASE: Involved in t(8;13)(p12;q12) chromosomal translocation
 CC which involves FGFR1 and ZNF198. The resulting transcript is a
 CC possible candidate for stem cell leukemia lymphoma syndrome/SCLL.
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- DATABASE: NAME-Atlas Genet. Cytoenet. Oncol. Haematol.;
 CC WWW-"http://www.infobiogen.fr/services/chronocancer/Genes/FGFR113.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 CC EMBL; X51803; CAA36101.1; -
 CC EMBL; X52833; CAA37015.1; -
 DR

DR EMBL: X66945: CAA47375.1; -
 DR EMBL: Y00665: CAA68679.1; -
 DR EMBL: M37122: AAA75007.1; -
 DR EMBL: M60485: AAA35845.1; -
 DR EMBL: M63887: AAA35958.1; -
 DR EMBL: M34185: AAA35836.1; -
 DR EMBL: M34186: AAA35837.1; -
 DR EMBL: X57118: CAA40400.1; ALT_TERM.
 DR EMBL: X57119: CAA40401.1; -
 DR EMBL: X57120: CAA40402.1; -
 DR EMBL: X57121: CAA40403.1; -
 DR EMBL: X57122: CAA40404.1; -
 DR EMBL: M34641: AAA35835.1; -
 DR EMBL: A29216: CAA01958.1; -
 DR PIR: S11692: TVHUG.
 DR PIR: S25420: S25420.
 DR PIR: S26739: S26739.
 DR PDB: 1FGK: 08-APR-97.
 DR PDB: 1AGW: 25-MAR-98.
 DR Genew: HGNC:3688: FGFR1.
 DR MIM: 101600: -
 DR InterPro: IPR000719: Euk_Pkinase.
 DR InterPro: IPR003006: Ig_Hc.
 DR InterPro: IPR003598: Ig_C2.
 DR InterPro: IPR001245: Tyr_Pkinase.
 DR Pfam: PF00047: Ig; 3.
 DR Pfam: PF00069: Pkinase; 1.
 DR PRINTS: PR00109: TYRKINASE.
 DR PRODOM: PD000001: Euk_Pkinase; 1.
 DR SMART: SM00408: IGC2; 3.
 DR SMART: SM00219: Tykrc; 1.

Query Match 6.9%; Score 141.5; DB 1; Length 822;
 Best Local Similarity 24.4%; Pred. No. 0.7;
 Matches 76; Conservative 41; Mismatches 93; Indels 101; Gaps 16;

QY 1 MMSWCKLLFMAVLTATCTAGAAVT-----PSVIDNLT----- 34
 DB 1 MMSWCKLLFMAVLTATCTARPSPTLPEDQAPGAPVEVSFLVHPGDLTLQRLRLDD 60
 QY 35 -----SKTE-NEYEVLYAKP-AGVYRLKQLKQLENVWF-FNVAVKD 73
 DB 61 VQSIINMLRDGVQALAESNRTRTGEEVQDSVPADSGLYACVTSSPGSDTTFYSVAVSD 120
 QY 74 ILNSRFKRNFKNVLESDLIPYKDLTSSNVVADPYKFLNKKERKFL----- 122
 DB 121 ALPSEDDDDDDSSSE---KETDNRKPNMPVAPF-WTSPKMEKKLHVPAAKTYKF 176
 QY 123 -----SYNTIKSIDTDINFAND-VLGTYKIILSEKYSKSLDSIKKIYNDKQEGNE 172
 DB 177 KCPSSGPNPTLRLKNGKE---FKPDHRIIGYKVRVATWSIIMDSV---VPSDKGN-- 227
 QY 173 KYLPFLNNIETLYTVNKKIDLFYIH-----LEAKV-LNYTYE-KSNVEVKIK----- 218
 DB 228 -----YTCIVNEVYSINHYYQLDVVERSPHPIQLAQLPANKTYALSSNVEPKYVSDP 263
 QY 219 --ELNYIKTIO 227
 DB 284 QPHIOMLKHE 294

RESULT 10
 RASO_METJA
 ID RASO_METJA STANDARD; PRT; 1005 AA.
 AC OS8718;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE DNA double-strand break repair rad50 Atpase.
 GN RAD50 OR Mj1322.
 OS Methanococcus jannaschii.

CC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 CC NCBI_TaxID:2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-963799; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Melnick R.K., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
 RA Uitterlidge T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).
 RL
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (by
 CC similarity).
 CC -1- SUBUNIT: Forms a complex with mre11 (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U67572: AAB99331.1; -
 DR TIGR: MJ1322: -
 DR InterPro: IPR003439: ABC_transportr.
 DR InterPro: IPR003405: SMC_C.
 DR InterPro: IPR003395: SMC_N.
 DR Pfam: PF02463: SMC_N; 1.
 DR Pfam: PF02483: SMC_C; 1.
 DR PRODOM: PD000006: ABC_transportr; 1.
 DR Kf DNA_repair: Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 DR NP_BIND 32 39 ATP (BY SIMILARITY).
 FT DOMAIN 158 849 COILED COIL (POTENTIAL).
 FT SEQUENCE 1005 AA; 119387 MW; 9BBBA8173E78F3 CRC64;

Query Match 6.9%; Score 141.5; DB 1; Length 1005;
 Best Local Similarity 22.4%; Pred. No. 0.87;
 Matches 86; Conservative 51; Mismatches 114; Indels 133; Gaps 18;

QY 34 LSKIEEN-EEVLYLKPLAGYRSIKROLNNVTFVNVK-----DILNSRFKRE 83
 DB 179 LERIEGLNFKNENKELKNNMSQL- -EKNKRLMEINDKLNKIKKEFDIELKFEWE 235
 QY 84 NFKNVLESDLIPYKDLTSSNVVADPYKFLN-KRDKFLSSVYITDSITDINF---A 139
 DB 236 NKKLLYE-----KFLNKEERRALELNKQELKILEYDINTVEEA 275
 QY 140 NDVLYGYKIILSEKYSKSLDSIKK-----YIN-DKQ-----GENKYLPELN 179
 DB 276 RETLNHRKDEYEKKSVDDELRIKESRLRELKSHYEDLVLTQLEIKIDIEKLEFIN 335
 QY 180 -----NIETL-----YKTVNKKIDLF----- 195
 DB 336 KSKYRDIDINDLTLNKKIKKEIERVETIKQLBELKLNBEIKIEYKRIKCECKRYE 395
 QY 196 -VHLEAKVINY-----YKSNVEYKIKELNYLKTIDKLADFKNNNFVGIAD 244
 DB 396 KYLELEKKAVEYNLTLEYITLLQEKKSIE---KNINDLETRINKLLEFRKNIDIESI-- 450
 QY 245 LSTDYNNNNLTLKFLSGVFNENLAKTYVL-NLIDGNLQGMN-----ISOHCYKRCOP 298

Db 451 -----ENSIKKEIEKKKLTENIOMKKEIKLNKKKLGINSIEIKRKLIDELKEVEGKCP 503

Oy 299 QNSGCFRHLDERECCLLYNKOE 322

Db 504 L---CKTPIDENKKMELINOHKTO 524

RESULT 11

FGRL MOUSE

ID FGRL MOUSE STANDARD: PRT: 822 AA.

AC P16092; 001736; Q61562;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Basic fibroblast growth factor receptor 1 precursor (EC 2.7.1.112)

DE (FGFR-1) (bFGF-R) (MFR).

GN HGFRI OR F1G.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=90160373; Pubmed=1689490;

RA Reid H.H., Wilks A.F., Bernard O.;

RT "Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed in the developing mouse brain."

RL Proc. Natl. Acad. Sci. U.S.A. 87:1596-1600(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Brain;

RX MEDLINE=90265603; Pubmed=2161096;

RA Safiran A., Avivi A., Orr-Urtreger A., Neufeld G., Lonal P.;

RT "The murine flg gene encodes a receptor for fibroblast growth factor."

RL Oncogene 5:635-643(1990).

RN [3]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RX MEDLINE=91207411; Pubmed=1708247;

RA Kounhara H., Kasayama S., Saito H., Matsumoto K., Sato B.;

RT "Expression cDNA cloning of fibroblast growth factor (FGF) receptor in mouse breast cancer cells: a variant form in FGF-responsive transformed cells."

RL Biochem. Biophys. Res. Commun. 176:31-37(1991).

RN [4]

RP SEQUENCE FROM N.A. (SHORT FORM).

RX MEDLINE=90272715; Pubmed=2161540;

RA Mansukhani A., Moscatelli D., Talafico D., Levytska V., Basilico C.;

RT "A murine fibroblast growth factor (FGF) receptor expressed in CHO cells is activated by basic FGF and Kaposi FGF."

RL Proc. Natl. Acad. Sci. U.S.A. 87:4378-4382(1990).

RN [5]

RP SEQUENCE OF 1-15 FROM N.A., AND ALTERNATIVE SPLICING.

RX MEDLINE=95100926; Pubmed=7802632;

RA Harada T., Saito H., Kounhara H., Kurebayashi S., Kasayama S., Terakawa N., Kishimoto T., Sato B.;

RT "Murine fibroblast growth factor receptor 1 gene generates multiple messenger RNAs containing two open reading frames via alternative splicing."

RL Biochem. Biophys. Res. Commun. 205:1057-1063(1994).

CC -1- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (AFGF).

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A LONG FORM (SHOWN HERE), A VARIANT FORM AND A SHORT FORM ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR FAMILY.

CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M28998; AAA37290.1; -;

DR EMBL; X51893; CAA36175.1; -;

DR EMBL; M65053; AAA37620.1; -;

DR EMBL; M33760; AAA37622.1; -;

DR EMBL; S74765; AAB32845.1; ALT_SEQ.

DR PIR; A34849; TVMSRG.

DR HSP; P11362; ITRG.

DR MGD; MGI:95522; Fgfr1.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00047; Ig; 3.

DR Pfam; PF00069; pkinase; 1.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00408; IGC2; 3.

DR SMART; SM00219; TyrKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.

KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding; Transmembrane; Signal; Phosphorylation; Repeat; Alternative splicing.

KW Immunoglobulin domain; Repeat; Alternative splicing.

FT SIGNAL 1 21

FT CHAIN 22 822

FT FT

FT FT

FT DOMAIN 22 376

FT TRANSMEM 377 397

FT DOMAIN 398 822

FT DOMAIN 48 108

FT DOMAIN 171 237

FT DOMAIN 270 348

FT DOMAIN 478 767

FT NP_BIND 484 492

FT BINDING 514 514

FT ACT_SITE 623 623

FT MOD_RES 654 654

FT DISULFID 55 101

FT DISULFID 178 230

FT DISULFID 277 341

FT CARBOHYD 77 77

FT CARBOHYD 117 117

FT CARBOHYD 227 227

FT CARBOHYD 240 240

FT CARBOHYD 264 264

FT CARBOHYD 296 296

FT CARBOHYD 317 317

FT CARBOHYD 330 330

FT CARBOHYD 30 30

FT VARSPLIC 31 119

FT VARSPLIC 148 149

FT VARSPLIC 229 229

FT CONFLICT 256 258

FT CONFLICT 270 270

FT CONFLICT 387 387

FT CONFLICT 440 440

FT CONFLICT 508 508

FT CONFLICT 544 544

FT CONFLICT 756 756

FT CONFLICT 765 765

FT SEQUENCE 822 AA; 91980 MW; D5A46955FA60926B CRC64;

1. EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

IG-LIKE C2-TYPE DOMAIN 1.

IG-LIKE C2-TYPE DOMAIN 2.

IG-LIKE C2-TYPE DOMAIN 3.

PROTEIN KINASE.

ATP (BY SIMILARITY).

ATP (BY SIMILARITY).

BY SIMILARITY.

PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

POTENTIAL.

POTENTIAL.

POTENTIAL.

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

Q -> QGSSSWPLVAVAA (IN VARIANT FORM).

MISSING (IN SHORT ISOFORM).

T -> S (IN REF. 4).

T -> S (IN VARIANT FORM).

ILQ -> HPS (IN REF. 1 AND 3).

G -> A (IN REF. 4).

I -> M (IN REF. 3).

G -> A (IN REF. 2).

V -> L (IN REF. 3).

I -> M (IN REF. 4).

R -> H (IN REF. 4).

E -> D (IN REF. 4).

FT NON_TER 1 1
 FT NON_TER 1251 1251
 SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.7%; Score 137.5; DB 1; Length 1251;
 Best Local Similarity 22.3%; Pred. No. 1.8; Mismatches 112; Indels 73; Gaps 14;
 Matches 69; Conservative 55; Mismatches 112; Indels 73; Gaps 14;

QY 31 DNLSIKENEYEVLT--YLRKPLAGVYSRLKOLEN--NVMTFVNVNVDILNSRPNKRENR 86
 DB DSKLEKIKTDFESLNALKTLLEGVNALKASSDNEHVGSKSEPNPAL-SELEKEET-- 952
 QY 87 NYLESDDLIPYKDLTSSNNVVDKPYFLNKKRDKFLSSYNYKDSI-----DPDI--NF 138
 DB 953 -----DIDSLNFMALDE-----LLKGRKTCVSRKLIKIDVTKEISDTEILINTI 997
 QY 139 ANDVLGYKKILSEKRY--SDLSIKKYINDKOGENEKYLKP-----LNNIETLTK 186
 DB 998 EKNVATAYATYKNTEDYQDVLTINHEHNTQVSNHEPTNFDKSNKSSSEELTKAVTQSK 1057
 QY 187 TVNHRIDLEVIHLEAKVLYTYEKSNVYKIKEL-----NYLKTIDOKLADFKNNNFVGI 242
 DB 1058 TTSKLGVIIE-----VVENEMNTIESSAKEIEALYNELKNKTSLSNEIYQTSNEVRL 1112
 QY 243 ADLSTDYNNHNNLTFLSTGWPENLAKTVLSNL-----LDGNTQGMNLISQ 289
 DB 1113 QMKSNAD-----KYIDVSKIFNTVLDFOKSNIVYNGHSINNVDKLGKIQELIDADS 1166
 QY 290 H---OCYRK 295
 DB 1167 SFTLESIRK 1175

RESULT 14

Y237_BUCAI STANDARD; PRT: 617 AA.
 ID Y237_BUCAI
 AC P57331
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BU237 precursor.
 GN BU237.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RA MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS."
 RL Nature 407:81-86(2000).
 CC -1- SIMILARITY: BELONGS TO THE SURFACE ANTIGEN D15 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AP001118; BAB12952.1; -
 DR InterPro: IPR000184; Bac_surfag_D15.
 DR Pfam: PF01103; Bac_surface_Ag; 1.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 20
 FT CHAIN 21 617 POTENTIAL.
 SQ SEQUENCE 617 AA; 71807 MW; 9C91CA6D53BC27B8 CRC64;

Query Match 6.6%; Score 136.5; DB 1; Length 617;

Best Local Similarity 22.4%; Pred. No. 0.95;
 Matches 82; Conservative 49; Mismatches 104; Indels 131; Gaps 21;

QY 3 SWKCLLENAVLTATCTAGAAVTPSVIDNIISKIENEVEVLYLRPLAGVYSIRKOLEN 62
 DB 204 SM-----WMLSKST-----YSPKEINDDLEHLKNFY-----LSN 233
 QY 63 NVMTFNVNVR--DIINSRPN-----KRENFKNVLESDLIPYKDLTSSNYVK 107
 DB 234 GYYFNVNNTKQVDFQDKKQVQDITIGVSGKKYKLSNF--FVNGNLFPPQKLI-INLI 290
 QY 108 DPYKFLNKKRP-----KFLSSYNYIKDSIDTD-----INFANDVLYG--- 146
 DB 291 NNENFYNRDKIDIIYKTRFLSEGYVMTKVIYVQDIDHKKKTIALNFNIMKKRYEYK 350
 QY 147 -----KIISEKYSDDLSTIKKYINDKOGENEKYLFFLNIIETLYTVNHRIDL 194
 DB 351 RIYFTGNETQDVRLLRKIK--QMEGKTFNKLVELGAKL-----LEKT-RYFSDV 398
 QY 195 FVHLEAKVLYTYEKSNEVYKIKEL-----LNY-----LKTIDOKL----- 230
 DB 399 KIIQ--KLNYSDSNQIDITYQVKQGTGSIKNGLGYSQTSFNLAFSQDNITGSGNS 455
 QY 231 --ADFRKNNFVGLADSTD-----YNNHNLTKFLSTGMV--FENLAKTVLSNL-LDGN 280
 DB 456 LKVDYIK-NDYOKYIDISTSYPEFFYNADLNARFFYNDKYNFNISMIKNTYGFBSN 514
 QY 281 LOGMLN 286
 DB 515 LGFLIN 520

RESULT 15

ARP_PLAFA

ID ARP_PLAFA STANDARD; PRT: 537 AA.
 AC P04931;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Asparagine-rich protein (AG319) (ARP) (Fragment).
 DE Plasmodium falciparum.
 OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=86176787; PubMed=2421257;
 RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,
 RA Brown G.V., Anders R.F., Kemp D.J.;
 RT "An asparagine-rich protein from blood stages of Plasmodium
 RT falciparum shares determinants with sporozoites."
 RL Nucleic Acids Res. 14:3089-3102(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M24328; AAA29491.1; -
 DR PIR: A23770; A23770.
 DR Malaria.
 KW Non_Ter 537 537
 FT NON_TER 537 537
 SQ SEQUENCE 537 AA; 63000 MW; A6D76B15318CC239 CRC64;

Query Match 6.5%; Score 135; DB 1; Length 537;
 Best Local Similarity 17.9%; Pred. No. 0.98;
 Matches 68; Conservative 67; Mismatches 152; Indels 92; Gaps 14;

55 SLKOLENNVMTFNVNVDILNSRPNKRENFVLESDDLIPYKDLTS-----SNVYKDPY 110

```

Db      121 SSMKNTDNKJDTSTYNNMKGITINDNNMMDYLRLNI--NNINIEKGSARKKFFTYNNMKNL 178
OY      111 KFL-----NKEKRDKFLSSYN-----YIKSIDTDINFANDV 142
Db      179 KFTQNNNDNNNNINEDNNNNNNNNNNNGVFSNTQNNNNNNRNNNSINIKRNLNNNNNNNNM 238
OY      143 --LG-----YKILSEKYKSDIDSIKKYINDKÖGENEKYLPPLNNIETLYKTVN 189
Db      239 NKMGSODKQNSNNNFY--NMYNYQNRKNSMNNNNNNNNNNNNNNNNNNNNNNNNNNNN 296
OY      190 HKID-----LFVHLEAKVLYNTYEKSNVEYKIKELNYLKI 226
Db      297 HNNNNNNNNNNNNNNNNNNNNNSIDSDSPNYHAHVKSMNNNNNNNNESNT--ANDPQNFQTN 355
OY      227 QDKLADFEKNNNFYGIADLSTDYNNHNNLLTKFLSTGMYF-----ENLAKTVLSNLLD 278
Db      356 NDNMKRENNNNNNYGYDDNTYHVNNNTPTSTDFBRAVGYNNNYLNNNNNNNSAVYNNSSN 415
OY      279 GNLOGMLNISOHCVKQCPONSGCFRHLDERECCLLNTYKÖBGDKCVENPNPTCENN 338
Db      416 GN-----NMKNENSENKNVADNNDLNN--NKNNNNNINIMNESINNNNTLNNNNEYNNQNN 469
OY      339 GGCDAKCTEEDSGSNGK 357
Db      470 NE-DED-----DDWGEUGE 483

```

Search completed: June 11, 2003, 16:03:05
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 16:02:30 ; Search time 80 Seconds
(without alignments)
989.026 Million cell updates/sec

Title: US-09-500-376-8
Perfect score: 2063
Sequence: 1 MMSWKCLLFMAVLVATLCT.....KPDSPYFDGIFCSHHHHH 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	91.1	400	5	003999 plasmodium
2	1879	91.1	539	5	025984 plasmodium
3	1879	91.1	539	5	025966 plasmodium
4	1879	91.1	539	5	09TYG1 plasmodium
5	1879	91.1	539	5	025976 plasmodium
6	1876	90.9	539	5	025972 plasmodium
7	1876	90.9	539	5	025981 plasmodium
8	1874	90.8	376	5	09BMG8 plasmodium
9	1873	90.8	539	5	025971 plasmodium
10	1873	90.8	539	5	025973 plasmodium
11	1196.5	57.9	651	5	025924 plasmodium
12	1194.5	57.9	373	5	025924 plasmodium
13	1193.5	57.9	569	5	025969 plasmodium
14	1193.5	57.9	569	5	025974 plasmodium
15	1193.5	57.9	569	5	025975 plasmodium
16	1193.5	57.9	569	5	025977 plasmodium

17	1193.5	57.9	569	5	025979	025979 plasmodium
18	1193.5	57.9	1694	5	09NHX1	09nhx1 plasmodium
19	1193.5	57.9	1694	5	09PTX5	09ptx5 plasmodium
20	1193.5	57.9	1704	5	09TZT4	09tzt4 plasmodium
21	1193	57.8	372	5	025725	025725 plasmodium
22	1192.5	57.8	569	5	025983	025983 plasmodium
23	1191.5	57.8	373	5	025721	025721 plasmodium
24	1189	57.6	372	5	043997	043997 plasmodium
25	1188	57.6	372	5	025726	025726 plasmodium
26	1184.5	57.4	373	5	025722	025722 plasmodium
27	1183.5	57.4	373	5	025723	025723 plasmodium
28	1180.5	57.2	569	5	025967	025967 plasmodium
29	1175.5	57.0	373	5	043995	043995 plasmodium
30	1174.5	56.9	569	5	025970	025970 plasmodium
31	1174.5	56.9	569	5	025980	025980 plasmodium
32	1174.5	56.9	569	5	025982	025982 plasmodium
33	1174	56.9	372	5	025719	025719 plasmodium
34	1174	56.9	372	5	025720	025720 plasmodium
35	1174	56.9	570	5	09TYG2	09tyg2 plasmodium
36	1174	56.9	570	5	025968	025968 plasmodium
37	1173.5	56.9	373	5	025727	025727 plasmodium
38	1173.5	56.9	569	5	025978	025978 plasmodium
39	1173	56.9	372	5	025717	025717 plasmodium
40	1170.5	56.7	373	5	043996	043996 plasmodium
41	1168.5	56.6	373	5	025728	025728 plasmodium
42	1168	56.6	372	5	025718	025718 plasmodium
43	1163	56.4	1720	5	025922	025922 plasmodium
44	1159	56.2	652	5	025923	025923 plasmodium
45	1066	51.7	219	5	09UB87	09ub87 plasmodium

ALIGNMENTS

RESULT 1

ID	003999	PRELIMINARY:	PRT:	400 AA.
AC	003999:			
DT	01-NOV-1996 (TREMblrel. 01, Created)			
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)			
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)			
DE	Merozoite surface protein 1 precursor (Fragment).			
GN	MSPI.			
OS	Plasmodium falciparum.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5833;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Blackman M.J., Ling I.T., Nicholls S.C., Holder A.A.;			
RT	"Proteolytic processing of the Plasmodium falciparum merozoite surface protein-1 produces a membrane-bound fragment containing two epidermal growth factor-like domains."			
RL	Submitted (May-1991) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; M64681; AAA29709.1;			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001245; Tyr_PKinase.			
DR	Pfam: PF00008; EGF; 1.			
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR_1.			
KW	Signal.			
FT	NON_TER	1	1	
FT	SIGNAL	<1	24	POTENTIAL.
FT	CHAIN	25	286	MEROZOITE SURFACE PROTEIN 1.
FT	CHAIN	287	400	MEROZOITE SURFACE PROTEIN 1.
SO	SEQUENCE	400 AA;	537P075058626AC2 CRC64;	
QY	Query Match	91.1%;	Score 1879;	DB 5; Length 400;
DB	Best Local Similarity	99.4%;	Pred. No. 1.3e-86;	
	Matches 355;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
	22 GAATPSVINDIISKIENEVEVLYLKPLAGYRSKQLENNVTFFVNVKDIINSFKN 81			
	23 GEAVTPSVINDIISKIENEVEVLYLKPLAGYRSKQLENNVTFFVNVKDIINSFKN 82			

QY 82 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 141
DB 83 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 142
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 201
DB 143 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 202
QY 202 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 261
DB 203 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 262
QY 262 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 321
DB 263 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 322
QY 322 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 378
DB 323 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 379

RESULT 2

Q25984 PRELIMINARY; PRT; 539 AA.
AC 025984;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major mezozoite surface protein (Fragment).
GN MSPI.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93295445; PubMed-8515786;
RA Jongwutives S., Tanabe K., Kanbara H.;
RT "Sequence conservation in the C-terminal part of the precursor to the major mezozoite surface proteins (MSPI) of Plasmodium falciparum from field isolates."
RL Mol. Biochem. Parasitol. 59:95-100(1993).
DR EMBL; D13356; BAA02617.1; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001245; Tyr-kinase.
DR Pfam: PF00008; EGF; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR EGF-like domain; Mezozoite.
KW NON_TER
FT 1
SQ SEQUENCE 539 AA; 6115 MW; 3788015F3127CB9E CRC64;

Query Match 91.1%; Score 1879; DB 5; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.8e-86;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAAVTPSYIDNLSIKENEYEVLYKPLAGYRSIKKOLENNVMTFNNVNDIILNSRKN 81
DB 162 GEAATPSYIDNLSIKENEYEVLYKPLAGYRSIKKOLENNVMTFNNVNDIILNSRKN 221
QY 82 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 141
DB 222 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 281
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 201
DB 282 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 341
QY 202 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 261
DB 342 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 401
QY 262 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 321
DB 402 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 461

QY 322 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 378
DB 462 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 518

RESULT 3

Q25966 PRELIMINARY; PRT; 539 AA.
AC 025966;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major mezozoite surface protein (Fragment).
GN MSPI.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93295445; PubMed-8515786;
RA Jongwutives S., Tanabe K., Kanbara H.;
RT "Sequence conservation in the C-terminal part of the precursor to the major mezozoite surface proteins (MSPI) of Plasmodium falciparum from field isolates."
RL Mol. Biochem. Parasitol. 59:95-100(1993).
DR EMBL; D13357; BAA02618.1; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001245; Tyr-kinase.
DR Pfam: PF00008; EGF; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR EGF-like domain; Mezozoite.
KW NON_TER
FT 1
SQ SEQUENCE 539 AA; 61145 MW; 787DE90C1D0ACDE7 CRC64;

Query Match 91.1%; Score 1879; DB 5; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.8e-86;
Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAAVTPSYIDNLSIKENEYEVLYKPLAGYRSIKKOLENNVMTFNNVNDIILNSRKN 81
DB 162 GEAATPSYIDNLSIKENEYEVLYKPLAGYRSIKKOLENNVMTFNNVNDIILNSRKN 221
QY 82 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 141
DB 222 RENEKNTVLESDLPKDLTSSNYVYKDPYKFLNKKRDKFLSSYNYIKDSIDTDINFAND 281
QY 142 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 201
DB 282 VLGYYKILSEKYSKSDLSIKKYINDKOGENEKYLPEFLNNIETLYKTVNHKIDLFVYHLEA 341
QY 202 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 261
DB 342 KVLNTYKESNVEYKIKELANTYKIQDKLADFKKNNNVGADLSTVDYNNHNLTKFLST 401
QY 262 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 321
DB 402 GAVFENLAKTYLSNLDGNLQGLNLSOHQCVKQCPONSGCFRHLDERECKCLLANTYKQ 461
QY 322 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 378
DB 462 EGDKCVENPNPTCNNGGCCADAKCTEEDSGSNCKKITTCCTKPDSTPLFDGIFCS 518
RESULT 4
Q25966 PRELIMINARY; PRT; 539 AA.
AC 025966;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major mezozoite surface protein (Fragment).
GN MSPI.

OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93295445; PubMed-8515786;
 RA Jongwutives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates.";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL: D13358; BAA02619.1; -
 DR Interpro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF, 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR, 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61075 MW; C70C2E10DEC4A101 CRC64;
 Query Match 91.1%; Score 1879; DB 5; Length 539;
 Best local Similarity 99.4%; Pred. No. 1.8e-86;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 22 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 81
 DB 162 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 221
 QY 82 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 141
 DB 222 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 281
 QY 142 VLGYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLANNIETLYKTVNKHIDLFVYHLEA 201
 DB 282 VLGYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLANNIETLYKTVNKHIDLFVYHLEA 341
 QY 202 KYLNTYKESNVEVRIKELNYLKTIDOKLADFRKNNNFVGIADLSTDYHNNILTKFLST 261
 DB 342 KYLNTYKESNVEVRIKELNYLKTIDOKLADFRKNNNFVGIADLSTDYHNNILTKFLST 401
 QY 262 GAVFENLAKTVLSNLDGMLQGLNLSOHCYKQKOPSGGFRHLDEBECKCLNTYKQ 321
 DB 402 GAVFENLAKTVLSNLDGMLQGLNLSOHCYKQKOPSGGFRHLDEBECKCLNTYKQ 461
 QY 322 EGDKCVENPPTCNENNCGCADAKCTEEDSGSNGKRITCECTKPDSPYLPFGIRCS 378
 DB 462 EGDKCVENPPTCNENNCGCADAKCTEEDSGSNGKRITCECTKPDSPYLPFGIRCS 518
 RESULT 5
 ID 025976 PRELIMINARY; PRT; 539 AA.
 AC 025976;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Major merozoite surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93295445; PubMed-8515786;
 RA Jongwutives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates.";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL: D13358; BAA02619.1; -
 DR Interpro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF, 1.
 DR PROSITE: PS00109; TYR-kinase.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61075 MW; C70C2E10DEC4A101 CRC64;

DR PROSITE: PS00109; PROTEIN_KINASE_TYR, 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61045 MW; 8BF9C64322E9A778 CRC64;
 Query Match 91.1%; Score 1879; DB 5; Length 539;
 Best local Similarity 99.4%; Pred. No. 1.8e-86;
 Matches 355; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 22 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 81
 DB 162 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 221
 QY 82 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 141
 DB 222 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 281
 QY 142 VLGYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLANNIETLYKTVNKHIDLFVYHLEA 201
 DB 282 VLGYKILSEKYSKSDLSIKKYINDKQGENEKYLPFLANNIETLYKTVNKHIDLFVYHLEA 341
 QY 202 KYLNTYKESNVEVRIKELNYLKTIDOKLADFRKNNNFVGIADLSTDYHNNILTKFLST 261
 DB 342 KYLNTYKESNVEVRIKELNYLKTIDOKLADFRKNNNFVGIADLSTDYHNNILTKFLST 401
 QY 262 GAVFENLAKTVLSNLDGMLQGLNLSOHCYKQKOPSGGFRHLDEBECKCLNTYKQ 321
 DB 402 GAVFENLAKTVLSNLDGMLQGLNLSOHCYKQKOPSGGFRHLDEBECKCLNTYKQ 461
 QY 322 EGDKCVENPPTCNENNCGCADAKCTEEDSGSNGKRITCECTKPDSPYLPFGIRCS 378
 DB 462 EGDKCVENPPTCNENNCGCADAKCTEEDSGSNGKRITCECTKPDSPYLPFGIRCS 518
 RESULT 6
 ID 025972 PRELIMINARY; PRT; 539 AA.
 AC 025972;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Major merozoite surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93295445; PubMed-8515786;
 RA Jongwutives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates.";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL: D13358; BAA02619.1; -
 DR Interpro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF, 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR, 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61076 MW; 72186F1412C57D1C CRC64;
 Query Match 90.9%; Score 1876; DB 5; Length 539;
 Best local Similarity 99.2%; Pred. No. 2.5e-86;
 Matches 354; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 22 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 81
 DB 162 GAATPSVIDNLSIKENEYEVLYIKPLAGYRSLSKKOLENNVMFPNVVKDILNSRFK 221
 QY 82 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 141
 DB 222 RENFKNVESDLIPYKDLTSSNYVADPKFLNKKERDFLSSYNIKISIDTDINFAND 281

Db 222 RENEKVLSDLPYKDLTSSNYYVADPYKFLNKKRDLKSLSYNTKIDSDTDINFAND 281
 QY 142 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEA 201
 Db 282 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEA 341
 QY 202 KVLNYYTEKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLST 261
 Db 342 KVLNYYTEKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLST 401
 QY 262 GMAVENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQ 321
 Db 402 GMAVENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQ 461
 QY 322 EGDGCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 378
 Db 462 EGDGCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 518

RESULT 7

ID 025981 PRELIMINARY; PRT; 539 AA.
 AC 025981;
 DT 01-NOV-1996 (TREMblrel, 01, Created)
 DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)
 DE Major merozoitte surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 CX NCBI_TaxID=5633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93295445; PubMed=8515786;
 RA Jongsavithes S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoitte surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates.";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL; D13363; BAA02624.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00008; EGF; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR EGF-like domain; Merzoitite.
 KW NON_TER 1
 FT 1
 SQ SEQUENCE 539 AA; 61046 MW; 3EED87473EE87B65 CRC64;

Query Match 90.9%; Score 1876; DB 5; Length 539;
 Best Local Similarity 99.2%; Pred. No. 2.5e-86;
 Matches 354; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 22 GAATPSYIDNLSKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFNVVVKDIINSRPNK 81
 Db 162 GEATTPSYIDNLSKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFNVVVKDIINSRPNK 221
 QY 82 RENEKVLSDLPYKDLTSSNYYVADPYKFLNKKRDLKSLSYNTKIDSDTDINFAND 141
 Db 222 RENEKVLSDLPYKDLTSSNYYVADPYKFLNKKRDLKSLSYNTKIDSDTDINFAND 281
 QY 142 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEA 201
 Db 282 VLGYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEA 341
 QY 202 KVLNYYTEKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLST 261
 Db 342 KVLNYYTEKSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLST 401
 QY 262 GMAVENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQ 321
 Db 402 GMAVENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQ 461
 QY 322 EGDGCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 378

Db 462 EGDGCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 518

RESULT 8

ID 09BMG8 PRELIMINARY; PRT; 376 AA.
 AC 09BMG8;
 DT 01-JUN-2001 (TREMblrel, 17, Created)
 DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
 DE Merzoitite surface protein 1 (Fragment).
 GN MSP-1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 CX NCBI_TaxID=5633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCB-1;
 RA Li X.R., Goel V.K., Liu S.C., Chishti A.H., Oh S.S.;
 RT "42 kDa subfragment of MSP-1 gene of Plasmodium falciparum FCB-1
 RT isolate.";
 RL EMBL; AF25919; AAK07641.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001245; Tyr-kinase.
 DR Pfam: PF00008; EGF; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
 KW NON_TER 1
 FT 1
 SQ SEQUENCE 376 AA; 43218 MW; EB2A72EDF231A5AF CRC64;

Query Match 90.8%; Score 1874; DB 5; Length 376;
 Best Local Similarity 99.7%; Pred. No. 2.1e-86;
 Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 AVTPSYIDNLSKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFNVVVKDIINSRPNK 83
 Db 1 AVTPSYIDNLSKIENEYEVLYLKLPLAGYRSLSKQLENNVMTFNVVVKDIINSRPNK 60
 QY 84 NFKNVLSDLPYKDLTSSNYYVADPYKFLNKKRDLKSLSYNTKIDSDTDINFANDYL 143
 Db 61 NFKNVLSDLPYKDLTSSNYYVADPYKFLNKKRDLKSLSYNTKIDSDTDINFANDYL 120
 QY 144 GYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEAKY 203
 Db 121 GYKILSEKYSKSDLSIKKYINDKOGENEKYLPLNNIETTLKYTNHKLIDLFVHLEAKY 180
 QY 204 LNTYTERSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLSTGM 263
 Db 181 LNTYTERSNVEYKIKELNLTAKTIODKLADFKKNNNFVGADISTDYNNHNNLTFRFLSTGM 240
 QY 264 VFENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQEG 323
 Db 241 VFENIAKTYVLSNLDGNLQGMNLISOHQCVKQCPQNSGCFRHLDERECCKLLNWKQEG 300
 QY 324 DKCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 378
 Db 301 DKCVENPNPTCENNGGCDADAKCTEEDSGSGNGKKTCTCPDSTYPLFDGIFCS 355
 RESULT 9
 ID 025971 PRELIMINARY; PRT; 539 AA.
 AC 025971;
 DT 01-NOV-1996 (TREMblrel, 01, Created)
 DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)
 DE Major merozoitte surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 CX NCBI_TaxID=5633;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-93295445; PubMed-8515786;
 RA Jongwutlives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL; D13359; BAA02620.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00008; EGF_1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61017 MW; 5CA9C651BB62B5E1 CRC64;

Query Match 90.8%; Score 1873; DB 5; Length 539;
 Best Local Similarity 99.2%; Pred. No. 3.5e-86;
 Matches 354; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 22 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 81
 DB 162 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 221
 QY 82 RENFKNVLESDLIPKDLTSSNYVVKDPYKFLNKKERKDFLSYNTIKDSITDINFAND 141
 DB 222 RENFKNVLESDLIPKDLTSSNYVVKDPYKFLNKKERKDFLSYNTIKDSITDINFAND 281
 QY 142 VLGYYKILSEKYSKSDLSIKKIYNDKOGENEKYLPLANNIEFLKYTVNKRIDLFTYHLEA 201
 DB 282 VLGYYKILSEKYSKSDLSIKKIYNDKOGENEKYLPLANNIEFLKYTVNKRIDLFTYHLEA 341
 QY 202 KYLNTYERKSNEVVKIKELNYLTKTIQDKLADFRKNNNFVGIADLSTDYNNHNLTKRFST 261
 DB 342 KYLNTYERKSNEVVKIKELNYLTKTIQDKLADFRKNNNFVGIADLSTDYNNHNLTKRFST 401
 QY 262 GAVFENLAKTVLSNLLDGMLOGLMNTSOHCYKQCPQNSGCFRHIDERECKCLLNTYKQ 321
 DB 402 GAVFENLAKTVLSNLLDGMLOGLMNTSOHCYKQCPQNSGCFRHIDERECKCLLNTYKQ 461
 QY 322 EBDKCVENPPTCENNGCGADAKCTEEDSGSNGKKITCECTKPDSTYPLFGIFCS 378
 DB 462 EBDKCVENPPTCENNGCGADAKCTEEDSGSNGKKITCECTKPDSTYPLFGIFCS 518

RESULT 10
 Q25973 PRELIMINARY; PRT; 539 AA.
 AC Q25973;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
 DE Major merozoite surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93295445; PubMed-8515786;
 RA Jongwutlives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates";
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL; D13359; BAA02622.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00008; EGF_1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER 1
 SQ SEQUENCE 539 AA; 61047 MW; 105C2E02974FB398 CRC64;

Query Match 90.8%; Score 1873; DB 5; Length 539;
 Best Local Similarity 99.2%; Pred. No. 3.5e-86;
 Matches 354; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 22 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 81
 DB 162 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 221
 QY 82 RENFKNVLESDLIPKDLTSSNYVVKDPYKFLNKKERKDFLSYNTIKDSITDINFAND 141
 DB 222 RENFKNVLESDLIPKDLTSSNYVVKDPYKFLNKKERKDFLSYNTIKDSITDINFAND 281
 QY 142 VLGYYKILSEKYSKSDLSIKKIYNDKOGENEKYLPLANNIEFLKYTVNKRIDLFTYHLEA 201
 DB 282 VLGYYKILSEKYSKSDLSIKKIYNDKOGENEKYLPLANNIEFLKYTVNKRIDLFTYHLEA 341
 QY 202 KYLNTYERKSNEVVKIKELNYLTKTIQDKLADFRKNNNFVGIADLSTDYNNHNLTKRFST 261
 DB 342 KYLNTYERKSNEVVKIKELNYLTKTIQDKLADFRKNNNFVGIADLSTDYNNHNLTKRFST 401
 QY 262 GAVFENLAKTVLSNLLDGMLOGLMNTSOHCYKQCPQNSGCFRHIDERECKCLLNTYKQ 321
 DB 402 GAVFENLAKTVLSNLLDGMLOGLMNTSOHCYKQCPQNSGCFRHIDERECKCLLNTYKQ 461
 QY 322 EBDKCVENPPTCENNGCGADAKCTEEDSGSNGKKITCECTKPDSTYPLFGIFCS 378
 DB 462 EBDKCVENPPTCENNGCGADAKCTEEDSGSNGKKITCECTKPDSTYPLFGIFCS 518

RESULT 11
 Q25924 PRELIMINARY; PRT; 651 AA.
 AC Q25924;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
 DE Merozoite surface antigen 1 (Fragment).
 OS Plasmodium falciparum.
 GN Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RO-71;
 RX MEDLINE-92275047; PubMed-1592091;
 RA Olafsson P., Matlie H., Certa U.;
 RT "Plasmodium falciparum: the repetitive MSA-1 surface protein of the
 RT RO-71 isolate is recognized by mouse antibody against the
 RT nonrepetitive repeat block of RO-33.";
 RL Exp. Parasitol. 74:381-389(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RO-71;
 RX MEDLINE-95354793; PubMed-7628566;
 RA Tolle R., Bujard H., Cooper J.A.;
 RT "Plasmodium falciparum: Variations within the C-terminal region of
 RT merozoite surface antigen-1.";
 RL Exp. Parasitol. 81:47-54(1995).
 DR EMBL; Z35329; CA84558.1; -;
 DR InterPro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF_1.
 FT NON_TER 1
 SQ SEQUENCE 651 AA; 74134 MW; AA2137B699255150 CRC64;

Query Match 58.0%; Score 1196.5; DB 5; Length 651;
 Best Local Similarity 59.8%; Pred. No. 2.3e-52;
 Matches 225; Conservative 56; Mismatches 74; Indels 21; Gaps 3;

QY 22 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 81
 DB 257 GAATPSPVIDNLSIKIENEYEVLYLKLPLAGVYRSLSKKOLENNVMTFNNVVKDILNSRFNK 315
 QY 82 RENFKNVLESDLIPKDLTSSNYVVKDPYKFLNKKERKDFLSYNTIKDSITDINFAND 141

```

DB 316 RKFLDVLESDLMQFKHISNEYIIEDSFKLINSQKNTLLSKYKIESYENDIKFAOE 375
OY 142 VLGYKILSEKYSKSDLSIKKYI-----NDKGENEYLPPLNNIE 182
DB 376 GISYEEKYAKKDDLESIKKYEKEKEFPSSPTTPSPAKTDEQKESKFLPLTNIE 435
OY 183 TLVKTVMHKIDLFVHLEAKVLTNTYKESNVEKIKELNTLKTODKLADPKKNNNFYGI 242
DB 436 TLNNLVNKKIDDYLLNFKAKINDCNVEDEAHVKTITKLSDKAIDDKIDLFKNHNDFEAI 495
OY 243 ADLSTDYNNHNLTKFLSTGAVFENLAKTVLSNLLDGLMGLNLSOHCYKQCPQNSG 302
DB 496 KKLINDTKKMDLKLSTGLV-QNFPNTIISKLEGFQDMLNLSOHCYKQCPQNSG 554
OY 303 CFRHLDERECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICE 362
DB 555 CFRHLDERECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICE 614
OY 363 CTRKPSYPLFDGIFCS 378
DB 615 CTRKPSYPLFDGIFCS 630

```

RESULT 12

```

ID 025724 PRELIMINARY; PRT; 373 AA.
AC 025724;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Merozoite surface protein 1, 42 kDa C-terminal region (Fragment).
GN MSP-1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=INFECTED HUMAN BLOOD SAMPLE;
RA Shi Y.-P., Alpers M.P., Poyoa M.M., Nahlen B.L., Olco A.G., Lai A.A.;
RT "Sequence of the C-terminal region of merozoite surface protein-1
RT (MSP-1) in field-derived Plasmodium falciparum.";
RL Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL, U20729; AAA62220.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
FT NON_TER 1
FT NON_TER 373
SQ SEQUENCE 373 AA; 42839 MW; A21A91B2076F415A CRC64;

```

Query Match 57.9%; Score 1194.5; DB 5; Length 373;
 Best Local Similarity 60.4%; Pred. No. 1.6e-52;
 Matches 223; Conservative 53; Mismatches 72; Indels 21; Gaps 2;

```

OY 30 IDNLSKRENEYEVLYLKLPLAGYRSLLKQLENNVMTFNNAVVKDILNSRFKREKRYVL 89
DB 6 MDNLSGFENEYDVLYLKLPLAGYRSLLKQLENNVMTFNNAVVKDILNSRLKRYFLDVL 65
OY 90 ESDILPYKDLTSSNVVDPYKFLNKEKRDKFLSSYNTIKSDITDINFADVGLYKYL 149
DB 66 ESDLMQFHHISSNEIIEDSFKLINSQKNTLLSKYKIESYENDIKFAOEGLSIYKIV 125
OY 150 SEKYSKSDLSIKKYI-----NDKGENEYLPPLNNIEFLYKTVN 189
DB 126 LAKRKDDLESIKKYIKEKEKEFPSSPTTPSPAKTDEQKESKFLPLTNIEFLYNNLV 185
OY 190 HKIDLPVHLEAKVLTNTYKESNVEKIKELNTLKTODKLADPKKNNNFVGIADLSTDY 249
DB 186 NKIDYLLNFKAKINDCNVEDEAHVKTITKLSDKAIDDKIDLFKNHNDFEAIKKLINDD 245
OY 250 NNNLTKFLSTGAVFENLAKTVLSNLLDGLMGLNLSOHCYKQCPQNSGCFRHLDE 309
DB 246 TKRMDLKLSTGLV-QNFPNTIISKLEGFQDMLNLSOHCYKQCPQNSGCFRHLDE 304

```

```

OY 310 REECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICECTKPSY 369
DB 305 REECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICECTKPSY 364
OY 370 PLFDGIFCS 378
DB 365 PLFDGIFCS 373

```

RESULT 13

```

ID 025969 PRELIMINARY; PRT; 569 AA.
AC 025969;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major merozoite surface protein (fragment).
GN MSP1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RE SEQUENCE FROM N.A.
RX MEDLINE-93295445; PubMed-8515786;
RA Jongwutlives S., Tanabe K., Kanbara H.;
RT "Sequence conservation in the C-terminal part of the precursor to the
RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
RT field isolates.";
RL Mol. Biochem. Parasitol. 59:95-100(1993).
DR EMBL, D13346; BAA02607.1; -.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 1.
KW EGF-like domain; Merozoite.
FT NON_TER 1
FT NON_TER 569
SQ SEQUENCE 569 AA; 64485 MW; A26C45914AF01396 CRC64;

```

Query Match 57.9%; Score 1193.5; DB 5; Length 569;
 Best Local Similarity 59.6%; Pred. No. 2.8e-52;
 Matches 224; Conservative 57; Mismatches 74; Indels 21; Gaps 3;

```

OY 22 GAATPSYIDNLSKRENEYEVLYLKLPLAGYRSLLKQLENNVMTFNNAVVKDILNSRFK 81
DB 175 GEALSVT-MDNLSGFENEYDVLYLKLPLAGYRSLLKQLENNVMTFNNAVVKDILNSRLK 233
OY 82 RENFNKYLESDLIPYKDLTSSNVVDPYKFLNKEKRDKFLSSYNTIKSDITDINFAND 141
DB 234 RKFLDVLESDLMQFKHISNEYIIEDSFKLINSQKNTLLSKYKIESYENDIKFAOE 293
OY 142 VLGYKILSEKYSKSDLSIKKYI-----NDKGENEYLPPLNNIE 182
DB 294 GISYEEKYAKKDDLESIKKYEKEKEFPSSPTTPSPAKTDEQKESKFLPLTNIE 353
OY 183 TLVKTVMHKIDLFVHLEAKVLTNTYKESNVEKIKELNTLKTODKLADPKKNNNFYGI 242
DB 354 TLNNLVNKKIDDYLLNFKAKINDCNVEDEAHVKTITKLSDKAIDDKIDLFKNHNDFEAI 413
OY 243 ADLSTDYNNHNLTKFLSTGAVFENLAKTVLSNLLDGLMGLNLSOHCYKQCPQNSG 302
DB 414 KKLINDTKKMDLKLSTGLV-QNFPNTIISKLEGFQDMLNLSOHCYKQCPQNSG 472
OY 303 CFRHLDERECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICE 362
DB 473 CFRHLDERECKCLNTKQEGDKCVENPNPTCNENNGCADACTEEDSGSNGKKTICE 532
OY 363 CTRKPSYPLFDGIFCS 378
DB 533 CTRKPSYPLFDGIFCS 548

```

RESULT 14

```

ID 025974 PRELIMINARY; PRT; 569 AA.

```

AC 025974;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE Major merozoite surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93295445; PubMed=8515786;
 RA Jongwutives S., Tanabe K., Kanbara H.;
 RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates."
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL: D13348; BAA02609.1.
 DR Interpro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF, 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER
 SQ SEQUENCE 569 AA; 64457 MW; A2691FE569E80396 CRC64;
 Query Match 57.9%; Score 1193.5; DB 5; Length 569;
 Best Local Similarity 59.6%; Pred. No. 2.8e-52;
 Matches 224; Conservative 57; Mismatches 74; Indels 21; Gaps 3;
 QY 22 GAATPSVYDNLTKLENEVEYLKPLAGVYSLSLKKOLENNMTNNVAVKDLNSRPNK 81
 DB 175 GEALSVT-MDNILSGENEVDVYLKPLAGVYSLSLKKOLENNMTNNVAVKDLNSRPNK 233
 QY 82 RENFNVLSDLPYKDLTSSNVVADPKFLNKRDRKFLSSYNTIKDSIDPTDINFAND 141
 DB 234 RKFFLDVLESDDLQGFHISSEYIIEDSRFLNSBQKNTLKSYYIKESVENDITFAQE 293
 QY 142 VLGYVILSEKYSKSDLSIKKYI-----NDKGENEKYLPFLNTE 182
 DB 294 GISYEVKYLAKYKDDLESIKKYI-----NDKGENEKYLPFLNTE 353
 QY 183 TLYKTVNHKIDLVHLEAKVLYNTEKSNVEYKIKELNYLKTIOKRLADFKNNPVG 242
 DB 354 TLYNNLVNKIDVLYLMLKAKINDCNVEKDEAHYKIKLSDLKAIIDKIDLFKNTDFEAI 413
 QY 243 ADLSTDYNNHNLTKFLSTGVFENLAKTVLSMLDGNLQGLNISQHCVCVKQCPONG 302
 DB 414 KRLINDTKKMDLGLSTGLV-QNFPNTIISKILGKRFQDMNLISQHCVCVKQCPONG 472
 QY 303 CERHLDREBECKCLLYKQBGKCVENPPTCENNGGCDADAKCTEEDSGSNGKKTICE 362
 DB 473 CERHLDREBECKCLLYKQBGKCVENPPTCENNGGCDADAKCTEEDSGSNGKKTICE 532
 QY 363 CTRPDSYPLFDGIFCS 378
 DB 533 CTRPDSYPLFDGIFCS 548

RESULT 15
 ID 025975 PRELIMINARY; PRT; 569 AA.
 AC 025975;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE Major merozoite surface protein (Fragment).
 GN MSP1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93295445; PubMed=8515786;
 RA Jongwutives S., Tanabe K., Kanbara H.;

RT "Sequence conservation in the C-terminal part of the precursor to the
 RT major merozoite surface proteins (MSP1) of Plasmodium falciparum from
 RT field isolates."
 RL Mol. Biochem. Parasitol. 59:95-100(1993).
 DR EMBL: D13349; BAA02610.1.
 DR Interpro: IPR000561; EGF-like.
 DR Pfam: PF00008; EGF, 1.
 KW EGF-like domain; Merozoite.
 FT NON_TER
 SQ SEQUENCE 569 AA; 64459 MW; ED4284B2867C9703 CRC64;
 Query Match 57.9%; Score 1193.5; DB 5; Length 569;
 Best Local Similarity 59.6%; Pred. No. 2.8e-52;
 Matches 224; Conservative 57; Mismatches 74; Indels 21; Gaps 3;
 QY 22 GAATPSVYDNLTKLENEVEYLKPLAGVYSLSLKKOLENNMTNNVAVKDLNSRPNK 81
 DB 175 GEALSVT-MDNILSGENEVDVYLKPLAGVYSLSLKKOLENNMTNNVAVKDLNSRPNK 233
 QY 82 RENFNVLSDLPYKDLTSSNVVADPKFLNKRDRKFLSSYNTIKDSIDPTDINFAND 141
 DB 234 RKFFLDVLESDDLQGFHISSEYIIEDSRFLNSBQKNTLKSYYIKESVENDITFAQE 293
 QY 142 VLGYVILSEKYSKSDLSIKKYI-----NDKGENEKYLPFLNTE 182
 DB 294 GISYEVKYLAKYKDDLESIKKYI-----NDKGENEKYLPFLNTE 353
 QY 183 TLYKTVNHKIDLVHLEAKVLYNTEKSNVEYKIKELNYLKTIOKRLADFKNNPVG 242
 DB 354 TLYNNLVNKIDVLYLMLKAKINDCNVEKDEAHYKIKLSDLKAIIDKIDLFKNTDFEAI 413
 QY 243 ADLSTDYNNHNLTKFLSTGVFENLAKTVLSMLDGNLQGLNISQHCVCVKQCPONG 302
 DB 414 KRLINDTKKMDLGLSTGLV-QNFPNTIISKILGKRFQDMNLISQHCVCVKQCPONG 472
 QY 303 CERHLDREBECKCLLYKQBGKCVENPPTCENNGGCDADAKCTEEDSGSNGKKTICE 362
 DB 473 CERHLDREBECKCLLYKQBGKCVENPPTCENNGGCDADAKCTEEDSGSNGKKTICE 532
 QY 363 CTRPDSYPLFDGIFCS 378
 DB 533 CTRPDSYPLFDGIFCS 548

Search completed: June 11, 2003, 16:04:32
 Job time : 82 secs

THIS PAGE BLANK (USPTO)